REMARKS

Claims 4, 31-32, 37-45 have been withdrawn from consideration by the examiner as being directed to a non-elected invention. Claims 1-2, 6-14, 24, 30, 33 and 37-45 are cancelled without prejudice. Claims 3, 5, 15, 16, 18, 20, 25, 26, 28, 29, 32, 34-36 and 46-60 have been amended to more clearly define the invention. No new matter is added.

Specification

The specification has been amended to include SEQ ID Nos. in order to comply with 37 C.F.R. §1.825. The sequence listing has also been amended so that the SEQ ID NO:s listed in the specification correspond to those in the sequence listing. In particular, the promoter of the elected invention is Mpr1217 (SEQ ID NO:22). The sequence listing filed previously lists a short primer sequence for SEQ ID NO:22, and listed the sequence of MPR1217 under SEQ ID NO:37. The attached sequence listing now lists SEQ ID NO:22 as having the sequence of Mpr1217, and not SEQ ID NO:37. Similarly, the attached sequence listing now lists SEQ ID NO:21 as having the sequence of Mpr1216, and not SEQ ID NO:17. These corrections do not entail new matter because the sequence of the promoter Mpr1217 had been listed as SEQ ID NO:37, and because the specification discloses the exact steps used to derive the Mpr promoter 1217 from the high molecular weight glutenin gene (PrHMWG-Dx5 (SEQ ID NO:1)), which was well known to the public at the time the invention was made (Accession number X12928; see page 20 of the instant specification).

In addition, the sequence listed on page 42, lines 26-27, has been designated as SEQ ID NO:38. SEQ ID NO:38 has been newly added to the sequence listing. Accordingly, attached herewith is a copy of the "Sequence Listing" (sheets 1/18 through 18/18) in paper form for the above-identified patent application as required by 37 C.F.R. §1.821(c) and a copy of the Sequence Listing in computer readable form as required by 37 C.F.R. §1.821(e). As required by 37 C.F.R. §1.821(f), Applicant's Attorney hereby states that the content of the "Sequence Listing" in paper form and the computer readable form of the "Sequence Listing" are the same

and, as required by 37 C.F.R. §1.821(g), also states that the submission includes no new matter. Applicants request the entry of the Sequence Listing in the above-captioned patent application to replace the existing sequence listing.

The specification has also been amended to insert the phrase "What is claimed is" at the top of page 71.

Restriction/election

The elected invention is drawn to the MPr1217 promoter (SEQ ID NO:22). Applicant also respectfully requests rejoinder of the MPr1216 promoter (SEQ ID NO:21) and the MPr1128 promoter (SEQ ID NO:4) into the elected group because a search of (SEQ ID NO:4) would necessarily result in the identification of SEQ ID NO:21 and SEQ ID NO:22. Figure 1 shows that the MPr1217 promoter (SEQ ID NO:22), the MPr1216 promoter (SEQ ID NO:21), and the MPr1128 promoter (SEQ ID NO:4) comprise three, two or one, respectively, copies of a sequence comprising a G Box positioned upstream of an enhancer element. Applicant respectfully submits that these three promoters are structurally different aspects of the same invention, and would not impose an undue search burden on the Examiner. Accordingly, Applicant respectfully requests rejoinder of Group 4, claims 2, 3, 5, 8-30, 32, 34-36, 46-61 with respect to SEQ ID NO:4, and Group 21, claims 2, 3, 5, 8-30, 32, 34-36, 46-61 with respect to SEQ ID NO:21, to the examination of elected Group 22, claims 2, 3, 5, 8-30, 32, 34-36, 46-61 with respect to SEQ ID NO:22, and the linking claims 1, 6 and 7.

Claim Objections

Claims 5, 29, 32, 34-36 and 46-61 have been amended to delete SEQ ID Nos. that refer to non-elected subject matter.

The office action states that Claim 57 and its dependent claims are multiple dependent claims. However, it was not intended that claim 57 be a multiple dependent claim. The language of Claim 57 has been amended to clarify that it is an independent claim, and therefore is not a multiple dependent claim. Claims 59 and 61 each depend on only one independent claim (claim 57).

Claim Rejections – 35 USC 112

Claims 1, 3, 5-6, 8, 10, 12-14, 16-18, 20, 23-26, 29-30, 32, 34-36 and 45-56 are asserted to be indefinite because they fail to begin with an article, such as "A" or "The". The pending claims have been amended so they each begin with an article.

Claim 5 is asserted to be indefinite because it recites "SEQ ID NO:8" twice. Both of the recitations of the phrase "SEQ ID NO:8" have been deleted from Claim 5.

Claims 5, 29, 32 and 34 are asserted to be indefinite because they recite multiple periods in the claims. Accordingly, the recitation of multiple periods in Claims 5, 29, 32 and 34 has been deleted.

Claim 46 is asserted to be indefinite because it recites the term "uence". Accordingly, the recitation of the term "uence" has been replaced with the term "sequence".

Claims 52-56, 58 and 60 are asserted to be indefinite because the claims refer to the method of claim 48 or 49, which are not methods. Accordingly, the recitation of "claim 49" has been deleted from claims 52-55, and the recitation of "claim 48" has been deleted from claim 56. The dependency of claims 58 and 60 have been changed from being dependent on claim 56 to being dependent on claim 57, the incorrect dependency being a result of typographical errors.

Claim Rejections – 35 USC 102

Claims 1-2, 6-28, 30, 35-36, 46-56, 58 and 60 are rejected under 35 U.S.C. 102(b) as being anticipated by Thomas et al.

Applicants submit that for a determination of anticipation to be proper, the prior art reference must disclose each and every limitation of the claim. *Atlas Powder Company et al. v. IRECO, Incorporated et al.*, 190 F.3d 1342, 1347 (Fed. Cir. 1999).

Applicant submits that the cited reference does not teach a chimeric promoter that comprises an AS2 box, and that accordingly the cited reference does not anticipate the instant claims which recite a chimeric promoter that comprise an AS2 box.

The office action describes the cited reference as teaching a chimeric promoter which "includes sequences from the CaMV promoter, which inherently contains an as1 (conferring root specific expression) and as2 box (conferring photosynthetic tissue expression)…" .

The Office action asserts that:

"An 'as1' and an 'as2' element are being interpreted as being present in any chimeric promoter that includes the CaMV promoter, as both of these sequences originate from the CaMV promoter".

Applicants respectfully disagree with said assertion, on the grounds that the chimeric promoter taught by Thomas et al. does not contain the as2 element.

Applicants first note that the chimeric promoter taught by Thomas et al. includes only a part of the CaMV promoter. Thomas et al teach that the CaMV "promoter carries a 5' deletion termination 89 bp upstream from the transcription initiation site and, therefore, retains the TATA and CCAAT-like boxes but not the putative enhancer sequences", see column 1 of page 1172.

Applicants also note that in the CaMV promoter, the as1 box extends from –82 to –64, while the as2 box is located at –85 to –105 (see page 1056, lines 1-15 of column 1 of an article by Verdaguer et al. (*Plant Molecular Biology* 37:1055-1067; 1998), Appendix A).

Because the sequences upstream of –89 of the CaMV promoter are not present in the referenced chimeric promoter, the as2 element which is located at –85 to –105, is also not present in the chimeric promoter taught by Thomas et al. Therefore, the chimeric promoter taught by Thomas et al. comprises a truncated CaMV promoter which contains the asp1 box, but not the asp2 box.

Because Thomas et al. does not teach a chimeric promoter comprising an as2 box, as required by claims 20-23, 25-28, Applicant respectfully requests reconsideration and withdrawal of the 102 rejection of claims 20-23 and 25-28.

Applicant contends that Thomas et al. does not teach a chimeric promoter with all the recited limitations of Claim 15, which is drawn to a chimeric promoter of gene expression comprising at least one transcriptional regulatory sequence from a gene encoding a high molecular weight wheat glutenin, which comprises a TATA box, a transcription start site, at least one enhancer box upstream of said TATA box and said transcription start site (+1), and at least one GATA box upstream of said at least one enhancer box, wherein said GATA box confers

light-regulatable expression on a transcription unit operably linked to said promoter. Specifically, the office action does not specify, nor can applicants find, where Thomas et al, teaches a chimeric promoter having a GATA box which confers light-regulatable expression on a transcription unit operably linked to said promoter. Accordingly, Applicant respectfully requests reconsideration and withdrawal of the 102 rejection of claims 15 .

Applicant contends that Thomas et al. does not teach a chimeric promoter with all the recited limitations of Claim 16, which is drawn to a chimeric promoter of gene expression comprising at least one transcriptional regulatory sequence from a gene encoding a high molecular weight wheat glutenin, which comprises a TATA box, a transcription start site, at least one enhancer box upstream of said TATA box and said transcription start site (+1), and comprising at least one cereal box upstream of the enhancer box. Applicant contends that the phrase "cereal box" is an art recognized term, which is defined as a motif containing a GCN4-like motif (GLM), (see column 2, page 711 of Guillaumie et al. Genome 47:705-713 (2004) Appendix "B"). Specifically, the office action does not specify where Thomas et al teaches a chimeric promoter which comprises a cereal box as required by claim 16. Further, the Office action does not specify where the Thomas et al. reference teaches (1) a chimeric promoter which comprises a cereal box that confers seed-specific expression on a transcription unit operably linked to the promoter, as required by claim 17; (2) a chimeric promoter which comprises two cereal boxes upstream of an enhancer box wherein no transcriptional regulatory sequences are between said cereal boxes, as required by claim 18; or (3) a chimeric promoter comprising two cereal boxes upstream of at least one enhancer box, wherein the cereal boxes are contiguous, as required by Claim 19. Accordingly Applicant respectfully requests reconsideration and withdrawal of the 102 rejection of claims 16-19.

Applicant contends that Thomas et al. does not teach a chimeric promoter, a cell comprising the promoter, nor a method for obtaining the cell, wherein the chimeric promoter contains all the recited limitations of the chimeric promoter recited in Claims 5, 29, 32, 34-36, 46-49, and 57-60. Claims 5, 29, 32, 34-36, 46-49, and 57-60 comprising a sequence selected from the group consisting of SEQ ID NO:22, SEQ ID NO:21 and SEQ ID NO:4, and these sequences are not taught by Thomas et al. Accordingly, Applicant respectfully requests

reconsideration and withdrawal of the 102 rejection of Claims 5, 29, 32, 34-36, 46-49, and 57-60.

Applicant contends that Thomas et al. does not teach a method for expressing a nucleic acid sequence encoding a polypeptide in a cell comprising a vector encoding a chimeric promoter comprising a sequence selected from the group consisting of SEQ ID NO:22, SEQ ID NO:21 and SEQ ID NO:4, as required by claim 50-56, because these sequences are not taught by Thomas et al. Accordingly, Applicant respectfully requests reconsideration and withdrawal of the 102 rejection of claims 50-56.

Claims Rejections – 35 USC 103

Claim 3 is rejected under USC 103(a) as being unpatentable over Thomas (as set forth in the rejection of claims 1, 2, 6-28, 30, 35, 36, 46-56, 58 and 60 under 35 USC(b)) in view of Halford et al.

To establish a *prima facie* case of obviousness, the Federal Circuit has stated in *In re Geiger* (815 F.2d 686, 688, 2 U.S.P.Q.2d 1276, 1278 (Fed. Cir. 1987) that "[o]bviousness cannot be established by combining the teachings of the prior art to produce the claimed invention, absent some teaching[,] suggestion or incentive supporting the combination."

Applicant traverses the rejection on the grounds that a teaching, suggestion or incentive supporting the combination of the cited references is absent in the 103 rejection.

Applicants agree with the Office Action's statement that Halford and Thomas both teach the expression of heterologous sequences using promoters derived from HMWG genes, and that Thomas et al. does not teach the use of the Glu-1Dx5 promoter. The Office action also states that Halford suggests "that the HMW regulatory sequences may be useful for driving high levels of expression [of foreign genes] in tobacco". However, Applicant notes that the Halford reference does not teach or suggest the use of chimeric promoters as required by claim 3.

The Office Action further states that "one of skill in the art would have been motivated to substitute the Glu-1Dx5 promoter taught by [Halford for the] Glu-D1-2 promoter taught by Thomas because Halford suggests using their particular promoters for the enhanced expression

of heterologous genes in tobacco, which is where Thomas is expressing their heterologous gene".

Applicant contends that the office action does not address where the Halford reference provides motivation for substituting the Glu-1Dx5 promoter taught by Halford for the Glu-D1-2 promoter taught by Thomas in a chimeric promoter. Applicant respectfully submits that the motivation does not provide for combining the references to produce a chimeric promoter as required by the instant claim, but only provides motivation for producing a heterologous gene. Because the office action does not supply motivation for combining the references to produce the chimeric promoter of the instant invention, no *prima facie* case of obviousness has been established. Applicant respectfully requests reconsideration and withdrawal of the rejection.

CONCLUSION

Applicant submits that all claims are allowable as written and respectfully request early favorable action by the Examiner. If the Examiner believes that a telephone conversation with Applicant's attorney/agent would expedite prosecution of this application, the Examiner is cordially invited to call the undersigned attorney/agent of record.

Respectfully submitted,

Date: September 9, 2004

Barbara Gyure

Name: Barbara Gyure
Registration No.: 34,614
Customer No.: 29933
Palmer & Dodge LLP
111 Huntington Avenue
Boston, MA 02199-7613
Tel: 617-239-0100

Colocation between a gene encoding the bZip factor SPA and an eQTL for a high-molecular-weight glutenin subunit in wheat (*Triticum aestivum*)

Sabine Guillaumie, Gilles Charmet, Laurent Linossier, Valérie Torney, Nathalie Robert, and Catherine Ravel

Abstract: The quality of wheat grain is largely determined by the quantity and composition of storage proteins (prolamins) and depends on mechanisms underlying the regulation of expression of prolamin genes. The endosperm-specific wheat basic region leucine zipper (bZIP) factor storage protein activator (SPA) is a positive regulator that binds to the promoter of a prolamin gene. The aim of this study was to map *SPA* (the gene encoding bZIP factor SPA) and genomic regions associated with quantitative variations of storage protein fractions using F_7 recombinant inbred lines (RILs) derived from a cross between *Triticum aestivum* 'Renan' and *T. aestivum* 'Récital'. *SPA* was mapped through RFLP using a cDNA probe and a specific single nucleotide polymorphism (SNP) marker. Storage protein fractions in the parents and RILs were quantified using capillary electrophoresis. Quantitative trait loci (QTLs) for protein were detected and mapped on six chromosome regions. One QTL, located on the long arm of chromosome 1B, explained 70% of the variation in quantity of the x subunit of *Glu-B1*. Genetic mapping suggested that *SPA* is located on chromosome arm 1L and is also present in the confidence interval of the corresponding QTL for Glu-B1x on 1BL, suggesting that *SPA* might be a candidate gene for this QTL.

Key words: Triticum aestivum, quantitative trait locus (QTL), single nucleotide polymorphism (SNP), storage protein activator (SPA), high-molecular-weight glutenin subunit (HMW-GS).

Résumé : La qualité du grain de blé est largement déterminée par la quantité et la composition en protéines de réserve (prolamines). Elle dépend également de mécanismes qui contrôlent la régulation de l'expression des gènes codant pour les prolamines. Le facteur de transcription bZIP, spécifique de l'albumen, activateur pour les protéines de réserve (SPA) est un régulateur du promoteur d'un gène de prolamine. SPA et les régions chromosomiques associées avec les variations quantitatives de différentes fractions protéiques ont été cartographiées à l'aide d'une population F_7 de lignées recombinantes obtenue à partir d'un croisement entre *Triticum aestivum* 'Renan' et *T. aestivum* 'Récital'. *SPA* a été d'abord assigné sur des lignées aneuploïdes grâce à une sonde d'ADNc. Il a aussi été cartographié génétiquement à l'aide d'un marqueur polymorphe basé sur une mutation ponctuelle (SNP). Les différentes fractions pour les protéines de réserve ont été mesurées par électrophorèse capillaire. Six loci à caractère quantitatif (QTLs) pour les protéines ont été détectés. Le QTL pour la quantité de gluténines de haut poids moléculaire (HMW-GS) codée par le locus *Glu-B1-1* explique 70 % de la variation phénotypique observée. L'assignement de SPA sur les lignées aneuploïdes révèle trois loci, situés sur le bras long des chromosomes homéologues du groupe 1. Le SNP développé montre qu'un des gènes de SPA co-localise avec le QTL pour la quantité de gluténines de haut poids moléculaire codée par le locus *Glu-B1-1*. *SPA* serait donc un gène candidat pour ce QTL.

Mots clés : Triticum aestivum, locus à caractère quantitatif (QTL), polymorphisme basé sur une mutation ponctuelle (SNP), activateur pour les protéines de réserve (SPA), gluténines de haut poids moléculaire (HMW-GS).

Received 4 November 2003. Accepted 8 March 2004. Published on the NRC Research Press Web site at http://genome.nrc.ca on 26 July 2004.

Corresponding Editor: F. Belzile.

S. Guillaumie, G. Charmet, V. Torney, N. Robert,[1] and C. Ravel.[2] Unité mixte de recherches INRA Université Blaise Pascal Amélioration et Santé des Plantes, 234 Avenue du Brézet, 63039 Clermont-Ferrand CEDEX 2, France.
L. Linossier. ULICE, Zone d'Activité Commericale des portes de Riom, 63200 Riom CEDEX, France.

[1]Present address: ISAB, rue Pierre Waguet, BP 30313, 60026 Beauvais CEDEX, France.
[2]Corresponding author (e-mail: ravel@clermont.inra.fr).

 doi: 10.1139/G04-031

Introduction

The endosperm of cereal grains is a storage organ containing proteins, oil, and carbohydrates, which are used by the embryo during germination. The most abundant wheat storage proteins exclusively synthesized in endosperm are prolamins, making up more than 90% of seed storage proteins. Wheat prolamins are divided into two groups: gliadins and glutenins. Gliadins consist of complex mixtures of monomeric proteins and are classified as α-, β-, γ-, and ω-gliadins according to their electrophoretic mobility, whereas glutenins are formed by high-molecular-weight and low-molecular-weight subunits (HMW-GS and LMW-GS), aggregated into high-mass polymers held together by intermolecular disulfide bonds. Gliadins and glutenins are important for bread-making quality, since they are involved in gluten network formation (Branlard and Dardevet 1985; Metakovsky et al. 1997; Branlard et al. 2001). Gliadins enhance gluten extensibility, whereas HMW-GS determine viscosity (strength) of the dough (Payne 1987; Shewry et al. 1992). In hexaploid wheat, HMW-GS are encoded by three loci, *Glu-A1*, *Glu-B1*, and *Glu-D1*, located on homeologous group-1 chromosomes (Payne 1987). Each locus is made of two tightly linked genes, e.g., *Glu-D1-1* and *Glu-D1-2*, which encode the x and y subunits, respectively. It should be noticed that *Glu-A1-2* (always) and *Glu-B1-2* (in some genotypes) have null alleles, so that Glu-B1y is generally absent.

The expression of seed storage protein genes is controlled by a network of regulatory proteins that interact with specific DNA elements and modulate their transcription. Transcriptional factors binding to specific DNA motifs are involved in such a regulation, as shown in the cases of Opaque-2 (Motto et al. 1989) and prolamin box binding factor (PBF) (Vicente-Carbajosa et al. 1997). In barley, four transcription factors involved in the expression of genes coding for hordein seed storage proteins have been identified (Mena et al. 1998; Vicente-Carbajosa et al. 1998; Oñate et al. 1999; Carbonero et al. 2000; Diaz et al. 2002). In wheat, Hammond-Kosack et al. (1993) demonstrated that two protein factors, endosperm box factors I and II (ESBF-I and ESBF-II), could bind to the bifactorial *cis*-acting element of an LMW-GS gene (Colot et al. 1989). This *cis*-acting element, called endosperm box, is generally well conserved in most seed storage protein gene promoters and is composed of the close juxtaposition of the GCN4-like motif (GLM) and the endosperm motif (EM), also termed the prolamin-box (P-box) (Colot et al. 1989). Albani et al. (1997) showed that a leucine-zipper protein called storage protein activator (SPA) was able to bind to the GLM and thus corresponded to ESBF-II, whereas Mena et al. (1998) reported that another transcriptional protein called wPBF (wheat prolamin-box binding factor), which would probably correspond to ESBF-I (Conlan et al. 1999), was able to bind EM. Albani et al. (1997) obtained the *SPA* cDNA from *Triticum aestivum* 'Chinese Spring', which is a 1647-bp-long sequence that contains an open reading frame coding for a protein of 409 amino acids. This cDNA is also characterized by several ATG codons in its leader sequence that determine short upstream open reading frames, a feature shared by genes encoding transcription factors, such as *Opaque-2* (Lohmer et al. 1993) and *Blz1* (Vicente-Carbajosa et al. 1998). They also demonstrated, by functional analysis, that SPA activates transcription through binding to the GLM of an LMW-GS gene of hexaploid wheat. This suggests that this gene is a target for SPA. No reports are available for any other targets of SPA or for its chromosomal location.

The aim of this study was to identify genomic regions associated with quantitative variations of storage protein fractions (QTLs for protein) and to map *SPA* using a population of RILs (F_7) derived from a cross between *Triticum aestivum* 'Renan' and *T. aestivum* 'Récital'.

Material and methods

Plant materials

Two French bread wheat cultivars, 'Renan', a superior bread-making wheat, and 'Récital', a standard bread-making wheat, were used to develop a mapping population. 'Renan' is characterized by a level of proteins that is higher than that in 'Récital' when grown in the same environment. These two cultivars have the same alleles for *Glu-A1-1*, *Glu-B1-2*, *Glu-D1-1*, and *Glu-D1-2* HMW (alleles 2*, 8, 5, and 10, respectively) and differ for the alleles of *Glu-B1-1* ('Renan' has allele 7 and 'Récital' has allele 6). A mapping population of 194 F_7 recombinant inbred lines (RILs) was obtained by single-seed descent (SSD) from the cross between these two lines. The RILs and the parental lines were sown in field trials at INRA plant breeding station of Clermont–Ferrand, France (45°47′ N, 3°10′ E, 329 m elevation) with two replications each in 1998 and 1999 as previously described (Groos et al. 2002). Plots were harvested in 1999 and 2000, respectively.

Aneuploid lines of 'Chinese Spring' used in this study consisted on 20 nullitetrasomic (NT) and 36 ditelosomic (Dt) lines kindly provided by Dr. S. Reader, John Innes Center, Norwich, UK.

Genomic DNA was extracted from fresh leaves. The extraction method reported by Lu et al. (1994) was used for RILs and parental lines, whereas the method previously described by Murigneux et al. (1993) was used for aneuploid lines.

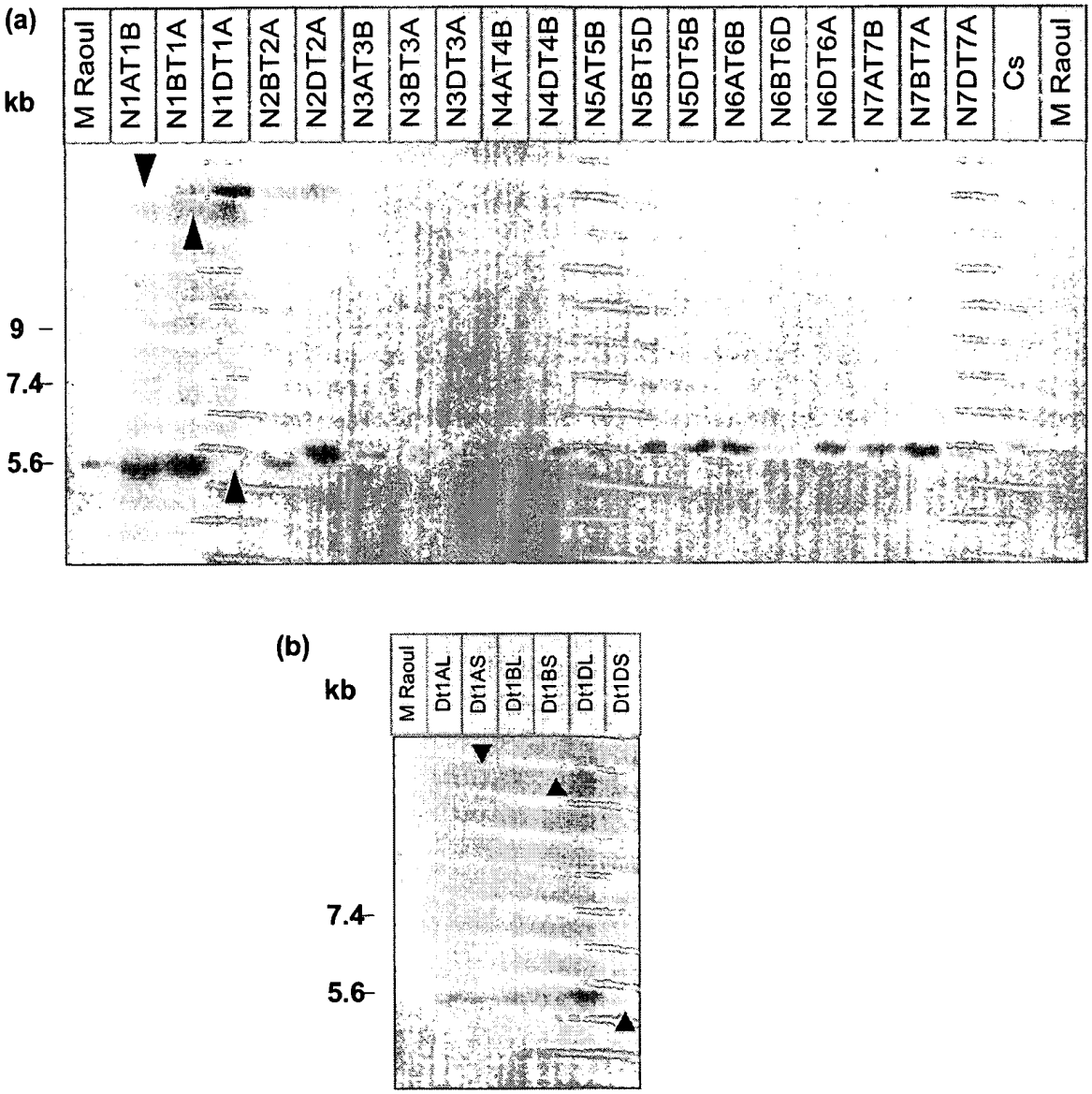
Genetic map

A genetic linkage map was obtained from this population and described in Groos et al. (2002). This map comprises 436 loci consisting of RFLP, microsatellites, and AFLP markers spread over the genome for a total length of 2740 cM, i.e., nearly 80% of the International Triticeae Mapping Initiative (ITMI) reference map total length (Leroy et al. 1997). This map also included some polymorphic storage protein loci, *Glu-B1-1* and *Gli-A1*. Linkage analyses were performed using Mapmaker/Exp version 3.06 (Lander et al. 1987). The Kosambi mapping function (Kosambi 1944) was applied to transform recombination frequencies into additive distances in centimorgans (cM). Linkage groups were assigned to chromosomes by comparison to reference maps using microsatellite loci (Cadalen et al. 1997).

QTL analysis

For each genotype, quantification of LMW glutenin subunits, HMW glutenin subunits, α+β+γ gliadins, and ω gliadins was achieved using capillary electrophoresis. Ex-

Fig. 1. Southern analysis. Chromosomal location of gene encoding SPA using *Triticum aestivum* 'Chinese Spring' (CS). (*a*) Nullitetrasomic lines (NT); T is followed by the name of the pair of duplicated chromosomes introduced in place of the lacking pair following N, e.g., N1AT1B corresponds to the nulli 1A tetra 1B line, i.e., chromosome 1A was substituted by chromosome 1B. (*b*) Ditelosomic lines (Dt); the lacking chromosome arm followed Dt, e.g., the ditelosomic line Dt1DS lacks the short arm of chromosome 1D. M Raoul, Raoul molecular marker. Disappearance of a band is marked by an arrow.



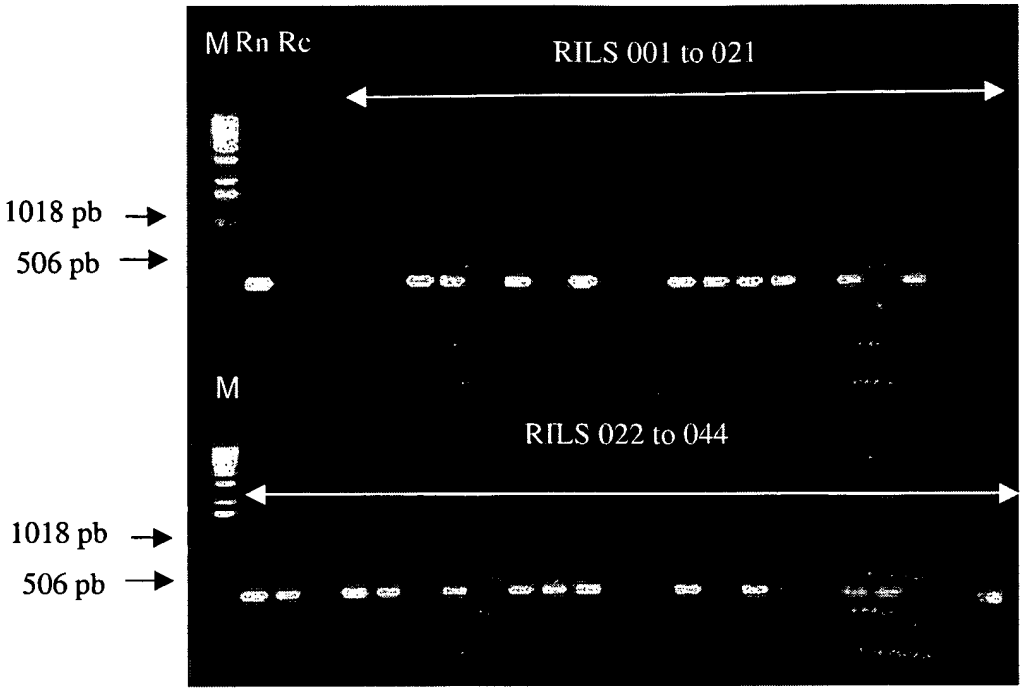
traction and separation procedures were adapted from those described elsewhere (Carceller and Aussenac 1999). Gliadins were specifically extracted in 50% *v/v* ethanol. Glutenins were precipitated with propanol-2 (initial concentration 50%; final concentration 70%), and then resuspended in the presence of SDS and 2-mercaptoethanol. All extracts were assayed by using a Lowry method (BioRad, Hercules, Calif., ref. No. 5000116) modified to be usable in the presence of reducing agent, i.e., after its neutralization by the addition of iodoacetamide. Glutenins were separated on the basis of their molecular mass for 15 min at 15 kV and 20 °C in a polyacrylamide gel (BioRad, ref. No. 148-5032) diluted by 5% *v/v* methanol (capillary gel electrophoresis). An uncoated silica capillary of 24 cm total length (18 cm for the separation) and 50 µm internal diameter (BioRad, ref. No. 148-3060) was used. Gliadins were separated in free zone (capillary zone electrophoresis), in a phosphate buffer (BioRad, ref. No. 148-5011), for 20 min at 22 kV and 45 °C. An uncoated silica capillary of 26 cm total length (20 cm for the separation) and 20 µm internal diameter (Thermo Finnigan, Ringoes, N.J. ref. TSP 20375) was used. Results were expressed as proportion of dry weight.

QTL detection was based on the Splus (1993) "homemade" program (Groos et al. 2002). Briefly, this program

Fig. 2. Alignment of sequences corresponding to the 5′ part of *SPA*. y09013 corresponds to 5′ part of the sequence Y09013. 'Récital' and 'Renan' are the corresponding sequences from genomic DNA of *Triticum aestivum* 'Récital' and *T. aestivum* 'Renan' cultivars, respectively. Each row has 60 nucleotides. The black box shows the start codon defined by Albani et al. (1997). Bold character indicates the SNP leading to a stop codon (TGA) instead of a TCA as in 'Renan'. The grey box shows an ATG codon that could constitute a potential start codon for 'Renan'.

```
y09013    TTCGCCGGTTCCCACGGGGTTAAAAAGTTTAAGCCATACCATTCCTCGCCCAATTTTGTT
Recital   TTCGCCGGTTCCCACGGGGTTAAAAAGTTTAAGCCATACCATTCCTCGCCCAATTTTGTT
Renan     TTCGCCGGTTCCCACGGGGTTAAAAAGTTTAAGCCATACCATTCCTCGCCCAATTTTGTT

y09013    CTAATCAATCACCTTTGCCCCAGTTCCATTTCGGAAAGCTTCCATGGAGCCCGTGTTCTT
Recital   CTAATCAATCACCTTTGCCCCAGTTCCATTTCGGAAAGCTTCCATGGAGCCCGTGTTCTT
Renan     CTAATCAATCACCTTTGCCCCAGTTCCATTTCGGAAAGCTTCCATGGAGCCCGTGTTCTT

y09013    CTCACTGGAGGAGGCGATGCCCGAGCCCGACTCTAACCCCTGCCGGACCTCGTCGCCGCC
Recital   CTCACTGGAGGAGGCGATGCCCGAGCCCGACTCTAACCCCTGCCGGACCTCGTCGCCGCC
Renan     CTGACTGGAGGAGGCGATGCCCGAGCCCGACTCTAACCCCTGCCGGACCTCGTCGCCGCC

y09013    GCTGGAGGCACACATGCTCGTCGCAGGACTCGGAGGAGTGGGCGCCGGCGAGGTCGTCGG
Recital   GCTGGAGGCACACATGCTCGTCGCAGGACTCGGAGGAGTGGGCGCCGGCGAGGTCGTCGG
Renan     GCTGGAGGCACACATGCTCGTCGCAGGACTCGGAGGAGTGGGCGCCGGCGAGGTCGTCGG

y09013    CGGGTGCGCGACGAACGAGTGCGCGACAGAATGGTGCTTCCAGAAGTTCGTGGACGAGCC
Recital   CGGGTGCGCGACGAACGAGTGCGCGACAGAATGGTGCTTCCAGAAGTTCGTGGACGAGCC
Renan     CGGGTGCGCGACGAACGAGTGCGCGACAGAATGGTGCTTCCAGAAGTTCGTGGACGAGCC

y09013    GTGGCTGCTCAACGTCCCCACCGCGCCAGTGGCGAACCCCGAAGCTTCGACGCTTTACCC
Recital   GTGGCTGCTCAACGTCCCCACCGCGCCAGTGGCGAACCCCGAAGCTTCGACGCTTTACCC
Renan     GTGGCTGCTCAACGTCCCCACCGCGCCAGTGGCGAACCCCGAAGCTTCGACGCTTTACCC

y09013    TAATCCCACGGCCGAGGGGAGCCGCAAGAGGCCGTACGACGTCCATGAGATGGTGGGCCC
Recital   TAATCCCACGGCCGAGGGGAGCCGCAAGAGGCCGTACGACGTCCATGAGATGGTGGGCCC
Renan     TAATCCCACGGCCGAGGGGAGCCGCAAGAGGCCGTACGACGTCCATGAGATGGTGGGCCC

y09013    GGAGGAGGTCATCCCCACGCCGCCTGCGGCGAGCCCGGTGGTGGACCCCGTGGCGTACAA
Recital   GGAGGAGGTCATCCCCACGCCGCCTGCGGCGAGCCCGGTGGTGGACCCCGTGGCGTACAA
Renan     GGAGGAGGTCATCCCCACGCCGCCTGCGGCGAGCCCGGTGGTGGACCCCGTGGCGTACAA

y09013    CGCGAT
Recital   CGCGAT
Renan     CGCGAT
```

Fig. 3. SNP polymorphism in a sample of F_7 recombinant inbreed lines (RILs) obtained from the cross between 'Renan' (Rn) and 'Récital' (Rc). This polymorphism was revealed by PCR using Pmut1 and Pmut2 primers. M, molecular weight marker. The number of each RIL used is indicated.



first performs a one-way analysis of variance (ANOVA) with every marker using a non-stringent threshold. A multiple regression is then carried out using the significant markers previously detected to select a subset of non-related significant markers. These markers, except those located on a given linkage group, are further used as covariates when scanning this chromosome using the marker regression method for fitting either a one-QTL (Kearsey and Hyne 1994) or two-QTL model (Hyne and Kearsey 1995). The 95% confidence intervals of QTL locations and additive effects were established by bootstrapping (Visscher et al. 1996) using 1000 replicates.

Sequence analysis

The cDNA sequence of *SPA* (Albani et al. 1997), accession number Y09013, was compared to sequences available in the Wheat Génoplante expressed sequence tag (EST) databank using BLASTN software (Altschul et al. 1997). This sequence was also used to design primers by Primer3 software (Rozen and Staletsky 2000) for PCR amplification from genomic DNA. Partial sequences of *SPA* obtained from each parental line were assembled and verified with the Staden package (Staden et al. 2000). In the case of doubtful peaks, an undefined nucleotide (N) was introduced into the sequence. Sequences obtained were aligned for mutation detection. Suitable pairs of primers characterizing detected mutation were designed with Primer3 software using a melting temperature of about 60 °C. In this pair, the last 3′ nucleotide of one primer corresponds to the mutation.

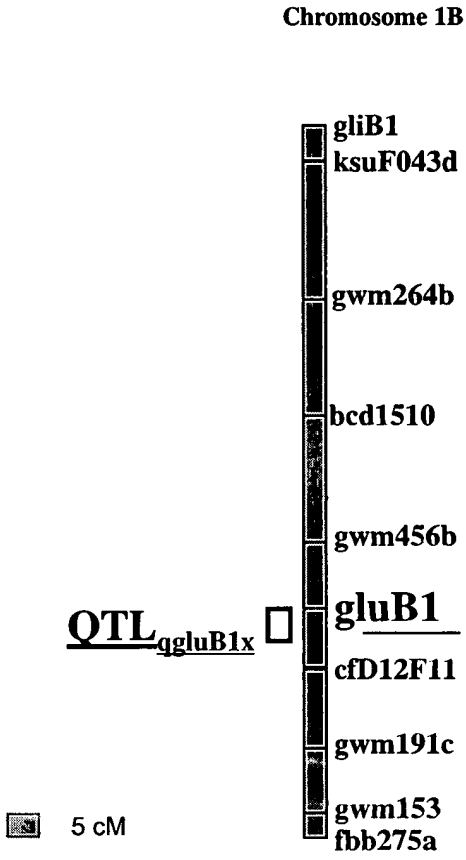
PCR primers and PCR amplification

Left and right primer sequences designed to amplify the 5′ segment of *SPA* were Pl (5′-GGCATCTCCGTCCTCTCTTC-3′) and Pr (5′-TCCAACTTCCGTCTGAGCAT-3′), respectively. PCRs were performed for each of the parental lines of the mapping population in a 40-μL final volume containing 80 ng of genomic DNA, 10% *v/v* glycerol, 250 μM of each dNTP, 0.4 μM of each primer, 1 U *Taq* DNA polymerase (Roche, Basel, Switzerland) and 1× *Taq* polymerase buffer. Cycles were 94 °C for 4 min; 35 cycles of 94 °C for 1 min, 60 °C for 1min 30 s, and 72 °C for 2 min ; and a final extension at 72 °C for 10 min. Amplification products were resolved on TAE 1× 1.2% *w/v* agarose gels. After running for 30 min at 100 V, gels were stained with ethidium bromide for 20 min in a 1 mg/L solution and visualized on a UV table. Double-strand sequencing of PCR products was performed by Genome Express. PCRs based on primers designed for SNP mutation were performed for RILs using the same protocol.

Southern blot analysis

Aneuploid lines were used to assign *SPA* to specific chromosomes and chromosome arms (Sears 1954, 1966; Anderson et al. 1992). Twelve micrograms of plant genomic DNA was digested for 5 h at 37 °C with *Hin*dIII, subjected to electrophoresis on a 0.8% w/v agarose gel, and transferred to a nylon membrane (Hybond-N+, Amersham, Piscataway, N.J.). A cDNA clone annotated as *SPA* and extracted from the Wheat Génoplante cDNA library was amplified using the

Fig. 4. Genetic mapping of *SPA* using the SNP revealed by Pmut1 and Pmut2 primers on chromosome 1B (112 cM in length). On the right, anchor markers (bold) and *SPA* marker, followed by the distance to the neighbouring anchor marker. The QTL for the quantity of x subunit of *Glu-B1* (QTL_{qgluBx}) is represented on left hand side, with its bootstrap confidence interval (white box).



SP6/T7 primers and the purified PCR product was radiolabeled by a random primer labelling kit (Megaprime, Amersham Biosciences). The blots were hybridized overnight at 65 °C with the SPA cDNA probe. After hybridization, the blots were washed at 65 °C once with a 2% *w/v* SSC solution for 20 min, followed by two washes with a 0.5× SSC and 0.1× SDS solution for 20 min. The blots were exposed to X-ray film to produce autoradiograms.

Results

BLAST analysis and RFLPs

BLAST analysis revealed that the Wheat Génoplante EST data contains only one contig, G468.114N24F010821.2.1, which is 1645 bp long and corresponds to the sequence Y09013 (1647 bp). The alignment between these two sequences showed a score of 2940 (*E* value = 0). The length of the alignment between these two sequences was 1645 bases, showing that there was 97% identity between them. The next alignment found by BLASTN showed only a score of 56 (*E* value = 1×10^{-6}) for a clone 1126 bp in length.

The contig G468.114N24F010821.2.1 contains seven ESTs, one of them (G468.114N24F010821) belonging to Génoplante. This EST, 585 bp long, belongs to a clone extracted from a cDNA library isolated from developing grains of 'Récital' and corresponds to the 3′ part of the sequence Y09013. Some differences existed between the sequence of the clone belonging to Génoplante and sequence Y09013. This may indicate that they do not characterize the same homoeologous gene. The cDNA probe used resolved a distinct RFLP locus on chromosome *1D*. Furthermore, fainter bands allowed detection of not so distinct RFLP loci on chromosomes *1A* and *1B*. These results suggest that at least one *SPA* locus is present per diploid genome (Fig. 1*a*). This assignment was confirmed by the results obtained with Dt lines, which also showed that *SPA* is located on the long arm of chromosomes 1A, 1B, and 1D (Fig. 1*b*).

Detection of a SNP and mapping of *SPA*

The primers Pl and Pr were positioned on the sense strand of the sequence Y09013 at 57 bp and 571 bp, respectively. These primers gave PCR products of about 500 bp from genomic DNA of 'Renan' and 'Récital'. This size fit that expected from the cDNA Y09013, indicating the absence of any intron in this part of the gene. Figure 2 shows the alignment between the sequence Y09013 and those of 'Renan' and 'Récital', confirming the absence of an intron in this part of the gene. Only one obvious difference appeared between the two cultivars at position 123, 17 bases after the potential start codon (Albani et al. 1997). The TCA codon present in 'Récital' and 'Chinese Spring' was substituted by TGA in 'Renan'. This mutation introduced a stop codon in the 'Renan' sequence, possibly leading to a truncated protein. Specific primers for each allele were designed. The forward and reverse primers used to map this mutation were Pmut1 (5′-ATGGAGCCCGTGTTCTTCTG-3′), corresponding to the 'Renan' allele, and Pmut2 (5′-CACCATCTCATGGACGTCGTA-3′), respectively, producing an amplicon of 311 bp (Fig. 3). This pair of primers revealed the expected band of 311 bp characterizing the 'Renan' allele in 84 RILs, whereas 109 RILs give no band, thus indicating that they had the 'Récital' allele. Thanks to this SNP, we were able to map *SPA* on the long arm of chromosome 1B, 1.3 cM away from the *Glu-B1-1* structural gene (Fig. 4). This result is in agreement with that obtained by Southern assignment.

QTL detection

QTLs for storage protein fractions were detected and mapped on six chromosome regions (Table 1). Three QTLs had small effects and large confidence intervals (on chromosomes 1A, 2B, and 5A), but the QTL on chromosome 7A appeared to have a relatively large effect, although its confidence interval was also quite large. One of the QTLs located on chromosome 1A colocated with its structural gene, *Gli-A1*. The QTL located on the long arm of chromosome 1B had a strong effect (r^2 = 70%) on the quantity of the x subunit of *Glu-B1* (Glu-B1x) and smaller effects on the quantity of y subunits of *Glu-D1* (Glu-D1y) and total HMW-GS, with the allele from 'Renan' giving higher Glu-B1x and total HMW-GS content, whereas it had a small, opposite effect on the quantity of Glu-D1y. This QTL explained 70% of phenotypic variation for Glu-B1x, and was located with

Table 1. Summary of QTLs identified in the 'Renan' × 'Récital' RIL population for quantitative variation of storage protein fractions measured by capillary electrophoresis.

Storage protein quantity	Chromosome	Field trial[a]	h2[b]	Detection power[c]	Map location (C.I.)[d]	Allele[e]
ω gliadins	1A	CF00	8.1	0.905	36.7 (6.9–69)	Récital
HMW-GS	1B	CF00	8.2	0.78	79.3 (16.2–85.2)	Renan
Glu-B1x	1B	CF99-CF00	70.3	1	80.8 (77.7–82.3)	Renan
Glu-D1x	1B	CF99	10.5	0.895	86.6 (66.5–96.7)	Renan
Glu-D1y	1B	CF99-CF00	42.2	1	78.1 (72.8–82.1)	Récital
α + β + γ gliadins	2B	CF99	11.6	0.935	81.4 (51.7–115.2)	Récital
LMW-GS	2B	CF99-CF00	11.4	0.92	89.6 (80.1–125.8)	Récital
HMW-GS	5A	CF99	15.5	0.975	54.1 (18.7–87.9)	Renan
Glu-D1x	5A	CF99	8.6	0.83	55 (7.1–143.1)	Renan
LMW-GS	7A	CF99-CF00	36.6	0.885	142 (82.6–150.2)	Renan
LMW-GS	7D	CF99	8.4	0.845	62.1 (50.4–74.5)	Renan

[a]CF99 and CF00 correspond to field trials harvested at Clermont–Ferrand in 1999 and 2000, respectively.
[b]Heritability, i.e., the percentage of phenotypic variance explained by the QTL.
[c]Corresponds to the proportion of the bootstrap replicates in which the QTL was found to be significant (out of 1000 replicates).
[d]Confidence intervals of QTL locations are the 2.5 and 97.5 percentile of the empiric bootstrap distribution.
[e]Indicates which parental allele is associated with the highest quantity.

fairly good precision. Two markers are located within its 8-cM confidence interval and can thus be postulated as candidate genes: the *Glu-B1-1* gene itself and *SPA*, a transcription factor previously identified as an activator of *Glu-D1* which codes for an LMW-GS.

Discussion

Five QTLs have been found with a high confidence level (percentage of bootstrap replicates giving significant results) over two years (1999 and 2000) for quantitative variation of several wheat storage proteins. Two QTLs colocate with their structural genes (*Gli-A1* and *Glu-B1*), whereas the others do not. In this study, we focused on the QTL on chromosome 1B in particular, which shows both the largest effect and a colocation with two possible candidate genes: the structural gene *Glu-B1-1* itself and *SPA*, the gene encoding a putatively interacting transcription factor.

Before now, the chromosomal location of wheat *SPA* was unknown. Southern blot analysis revealed that one *SPA* locus exists on the long arm of each homoeologous group 1 chromosome. However, this technique does not permit determination of whether these loci contain a single or several tandem copies of *SPA*. An SNP marker was developed, which made it possible to locate this gene on chromosome 1B. This is in agreement with results obtained by mapping with RFLPs.

The SNP used to map *SPA* corresponds to the allele from 'Renan'. It introduced a stop codon into the coding sequence defined by Albani et al. (1997). This stop is fairly close to the start codon of the published sequence (it is the 6th codon after the start codon). It is followed by four more codons and another ATG codon, which could constitute a new potential start codon. This organisation creates a new short upstream open reading frame (uORF), in addition to those described by Albani et al. (1997). The presence of a few short uORFs is frequent in the mRNA of transcriptional factors where they constitute translational control elements. For example, translation of mRNA of the maize transcriptional activator Opaque-2 is inhibited by ORFs (Lohmer et al. 1993). Moreover, such an organisation may also lead to a shorter protein with a non-altered bZip domain than that predicted by Albani et al. (1997). Protein size polymorphism is also reported for maize bZip activator Opaque-2 (Hartings et al. 1995). Surprisingly, this truncated protein corresponds to 'Renan,' which has the greatest quantity of Glu-B1x.

The QTL for Glu-B1x had two candidate genes: *SPA* and *Glu-B1-1*, both located within the confidence interval of the QTL. This result indicates that both candidates may be involved in the synthesis of this HMW-GS.

SPA function has been elucidated (Albani et al. 1997; Conlan et al. 1999) and it was demonstrated that this bZip protein can activate transcription of an LMW-GS gene via the GLM. SPA may be involved in the activation of many other wheat prolamin genes, which generally contain conserved GLM in their promoter, since the promoters of HMW-GS genes are characterized by a motif called "cereal box", which contains a GLM. Results reported by Norre et al. (2002) supported a possible interaction between Opaque-2, a bZIP factor of maize, and DNA-binding sites of the HMW-GS promoter for *Glu-D1-1*. Moreover, group 1L chromosome arms often appear to have positive effects on HMW-GS genes (Wanous et al. 2003). Therefore, the location of *SPA* in the confidence interval of a QTL for the x unit of *Glu-B1-1* might suggest that SPA would also be able to activate expression of HMW-GS genes. SPA might have several targets. Such a phenomenon has already been demonstrated for Opaque-2. This transcriptional factor was shown to differentially influence transcription rates of α-zein genes, which shared common *cis*-acting elements, including the binding site for Opaque-2 (Kodrzycki et al. 1989). It is also required for the expression of a number of non storage protein genes including *b-70* and *b-32* (Di Fonzo et al. 1979; Lohmer et al. 1991).

As expected, the structural gene *Glu-B1-1* is also a candidate for the QTL on chromosome 1B. This result may be explained by polymorphism between the promoters of alleles 6 and 7 of *Glu-B1-1*. Considerable polymorphism was detected between the promoters of several HMW-GS genes by Anderson et al. (1998). These promoters can differ by the

cereal box copy number. More recently, a variability has been reported between two versions of the promoter of allele 7 of *Glu-B1-1*: one version presents a duplication of a 43-bp element, which is systematically associated with an overexpression of Bx7 (Juhàsz et al. 2003). However, the polymorphism of the promoters for *Glu-B1-1* is not yet fully described for every allele.

Therefore, our results show that it is difficult to separate the respective effects of the two candidates and that further analysis is necessary. It would be interesting to create and analyse new RILs with parents segregating for only one candidate gene. Ultimately, over expression and knock out of *SPA* could be of interest in studying the influence of this gene on storage protein synthesis. Moreover, it should be remembered that other transcriptional factors (Carbonero et al. 2000; Diaz et al. 2002) are involved in the complex regulatory pathway of cereal storage proteins, the dissection of which has just begun.

Acknowledgements

The authors wish to thank B. Debote and the staff of the experimental unit, INRA Clermont–Ferrand, for high-quality experimentation and Dr. F. Vear for helpful comments and English revisions. This research was funded by the French federative genomic program Génoplante.

References

Albani, D., Hammond-Kosack, M.C.U., Smith, C., Conlan, S., Colot, V., Holdsworth, M., and Bevan, M. 1997. The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. Plant Cell, **9**: 171–184.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Jinghui Zhang, Z.Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389–3402.

Anderson, J.A., Ogihara, Y., Sorrells, M.E., and Tanksley, S.D. 1992. Development of a chromosomal arm map for wheat based on RFLP markers. Theor. Appl. Genet. **83**: 1035–1043.

Anderson, O.D., Abraham-Pierce, F.A., and Tam, A. 1998. Conservation in wheat high-molecular-weight glutenin promoter sequences: comparisons among loci and among alleles of the *GLU-B1-1* locus. Theor. Appl. Genet. **96**: 568–576.

Branlard, G., and Dardevet, M. 1985. Diversity of grain protein and bread wheat quality. II. Correlation between high molecular weigh subunits of glutenin and flour quality characteristics. J. Cereal Sci. **3**: 345–354.

Branlard, G., Dardevet, M., Saccomano, R., Lagoutte, F., and Gourdon, J. 2001. Genetic diversity of wheat storage proteins and bread wheat quality. Euphytica, **119**: 59–67.

Cadalen, T., Bœuf, C., Bernard, S., and Bernard, M. 1997. An intervarietal molecular marker map in *Triticum aestivum* L. Em. Thell., and comparison with a map from wide cross. Theor. Appl. Genet. **94**: 367–377.

Carbonero, P., Vicente-Carbajosa, J., Mena, M., Oñate, L., Lara, P., and Diaz, I. 2000. BZip and Dof transcription factors in the regulation of gene expression in barley endosperm. *In* Seed biology: advances and applications. *Edited by* M. Black, K.J. Bradford, and J. Vazquez-Ramos. CABI Publishing, Wallingford, UK. pp. 27–41.

Carceller, J.L., and Aussenac, T. 1999. Accumulation and changes in molecular size distribution of polymeric proteins in developing grains of hexaploid wheats: role of the dessication phase. Aust. J. Plant Physiol. **26**: 301–310.

Colot, V., Bartels, D., Thompson, R., and Flavell, R. 1989. Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin genes. Mol. Gen. Genet. **216**: 81–90.

Conlan, R., Hammond-Kosack, M., and Bevan, M. 1999. Transcription activation mediated by the bZIP factor SPA on the endosperm box is modulated by ESBF-1 *in vitro*. Plant J. **19**: 173–181.

Diaz, I., Vicente-Carbajosa, J., Abraham, Z., Martinez, M., Isabel-Lamoneda, I., Carbonero, P. 2002. The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. Plant J. **29**: 453–464.

Di Fonzo, N., Formasari, E., Salamini, F., Reggiani, R., and Soave, C. 1979. Interaction of the mutants *floury-2*, *opaque-7* with *opaque-2* in the synthesis of endosperm proteins. J. Hered. **71**: 397–402.

Groos, C., Robert, N., Bervas, E., and Charmet, G. 2002. Analysis of genetic, environmental and genetic × environmental components for grain protein content, grain yield and thousand kernel weight in bread wheat. Theor. Appl. Genet. **104**: 39–47.

Hammond-Kosack, M.C.U., Holdsworth, M.J., and Bevan, M.W. 1993. *In vivo* footprinting of a low molecular weight glutenin gene (LMWG-1D1) in wheat endosperm. EMBO J. **12**: 545–554.

Hartings, H., Lazzaroni, N., Rossi, V., Riboldi, G.R., Thompson, R.D., Salamini, F., and Motto, M. 1995. Molecular analysis of *opaque-2* alleles from *Zea mays* L. reveals the nature of mutational events and the presence of a hypervariable region in the 5′ part of the gene. Genet. Res. **65**: 11–19.

Hyne, V., and Kearsey, M.J. 1995. QTL analysis: further uses of "marker regression". Theor. Appl. Genet. **91**: 471–476.

Juhàsz, A., Gàrdonyi, M., Tamàs, L., and Bedõ, Z. 2003. Characterization of the promoter region of Glu-1Bx7 gene from everexpressing lines of an old Hungarian wheat variety. *In* Proceedings of the 10th International Wheat Genetics Symposium, Paestum, Italy, 1–6 September 2003. *Edited by* N.E. Pogna, M. Romanò, E.A. Pogna, and G. Galterio. Instituto Sperimentale par la Cerealicoltura, Paestum, Italy. pp. 1348–1350.

Kearsey, M.J., and Hyne, V. 1994. QTL analysis: a simple "marker regression" approach. Theor. Appl. Genet. **89**: 698–702.

Kodrzycki, R., Boston, R.S., and Larkins, B.A. 1989. The *opaque-2* mutation of maize differentially reduced zein gene transcription. Plant Cell, **1**: 105–114.

Kosambi, D.D. 1944. The estimations of map distance from recombination value. Ann. Eugenics, **12**: 172–175.

Lander, E.S., Geen, P., Abrahanson, J., Barlow, A., Daly, M.J., Lincoln, S.E., and Newbur, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, **1**: 174–181.

Leroy, P., Nègre, S., Tixier, M.H., Perretant, M.R., Sourdille, P., et al. 1997. A genetic reference map for the bread wheat genome – *Triticum aestivum* L. em. Thell. *In* Progress in Genome Mapping of Wheat and Related Species. Joint Proceedings of 5th and 6th Public Workshops of the International Triticeae Mapping Initiative, Norwich, UK, 1–3 September 1995 and Sydney, Australia, 30–31 August 1996. Report No. 18. *Edited by* P.E. McGuire, and C.O. Qualset. University of California Genetic Resources Conservation Program, Davis, Calif. pp. 134–140.

Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F., and Thompson, R.D. 1991. The maize regulatory locus *Opaque-2* encodes a DNA-binding protein which activates the transcription of the *b-32* gene. EMBO J. **10**: 617–624.

Lohmer, S., Maddaloni, M., Motto, M., Salamini, F., and Thompson, R.D. 1993. Translation of the mRNA of the maize transcriptional activator Opaque-2 is inhibited by upstream open reading frames present in the leader sequence. Plant Cell, **5**: 65–73.

Lu, Y.H., Merlino, M., Isaac, P.G., Bernard, M., and Leroy, P. 1994. A comparative analysis between ^{32}P and digoxygenin-labelled single copy probes for RFLP detection in wheat. Agronomie, **14**: 33–39.

Mena, M., Vicente-Carbajosa, J., Schmidt, R.J., and Carbonero, P. 1998. An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from prolamin-box of a native B-hordein promoter in barley endosperm. Plant J. **16**: 53–62.

Metakovsky, E.V., Félix, I., and Branlard, G. 1997. Association between dough quality (W value) and certain gliadin alleles in French common wheat cultivars. J. Cereal Sci. **26**: 371–373.

Motto, M., Di Fonzo, N., Hartings, H., Maddaloni, M., Salamini, F., Soave, C., and Thompson, R.D. 1989. Regulatory genes affecting maize storage protein synthesis. Oxford Surv. Plant Mol. Cell. Biol. **6**: 87–114.

Murigneux, A., Barloy, D., Leroy, P., and Beckert, M. 1993. Molecular and morphological evaluation of doubled haploid lines in maize: 1. Homogeneity within DH lines. Theor. Appl. Genet. **86**: 837–842.

Norre, F., Peyrot, C., Garcia, C., Rancé, I., Drevet, J., Theisen, M., and Gruber, V. 2002. Powerful effect of an atypical bifactorial endosperm box from wheat HMWG-Dx5 promoter in maize endosperm. P. Mol. Biol. **50**: 699–712.

Oñate, L., Vicente-Carbajosa, J., Lara, P., Diaz, I., and Carbonero, P. 1999. Barley BLZ2, a seed-specific bZIP protein that interacts with BLZ1 *in vivo* and activates transcription from the GCN4-like motif of B-hordein promoters in barley endosperm. J. Biol. Chem. **274**: 9175–8182.

Payne, P.I. 1987. Genetics of wheat storage proteins and the effects of allelic variation on breadmaking quality. Annu. Rev. Plant Physiol. **38**: 141–153.

Rozen, S., and Staletsky, H.J. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132**: 365–386.

Sears, E.R. 1954. The aneuploid of common wheat. University of Missouri Res. Bull. **572**: 1–58.

Sears, E.R. 1966. Nullisomic-tetrasomic combinations in hexaploid wheat. *In* Chromosome manipulations and plant genetics. *Edited by* R. Riley, and K.R. Lewis. Olivier and Boyd, Edinburgh, UK. pp. 29–45.

Shewry, P.R., Halford, N.G., and Tatham, A.S. 1992. The high molecular weight subunits of wheat glutenin. J. Cereal Sci. **15**: 105–120.

Splus. 1993. Splus guide to statistical and mathematical analyses. Mathsoft, Seattle, Wash.

Staden, R., Beal, K.F., and Bonfield, J.K. 2000. The Saden package, 1998. Methods Mol. Biol. **132**: 115–130.

Vicente-Carbajosa, J., Moose, S.P., Parsons, R.L., and Schmidt, R.J. 1997. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc. Natl. Acad. Sci. U.S.A. **94**: 7695–7690.

Vicente-Carbajosa, J., Oñate, L., Lara, P., Diaz, I., and Carbonero, P. 1998. Barley BLZ1 : a bZip transcriptional activator that interacts with endosperm-specific gene promoters. Plant J. **13**: 629–640.

Visscher, P.M., Thompson, R., and Haley, C.S. 1996. Confidence intervals in QTL mapping by bootstrapping. Genetics, **143**: 1013–1020.

Wanous, M.K., Munkvold, J.D., Kruse, J.D., Brachman, E.E., Klawiter, M.A., and Fuehrer, K.J. 2003. Identification of chromosome arms influencing expression of the HMW glutenins in wheat. Theor. Appl. Biol. **106**: 213–220.

Plant Molecular Biology **37**: 1055–1067, 1998.

Functional organization of the cassava vein mosaic virus (CsVMV) promoter

Bertrand Verdaguer[1], Alexandre de Kochko, Charles I. Fux, Roger N. Beachy and Claude Fauquet*

*International Laboratory for Tropical Agricultural Biotechnology (ILTAB/ORSTOM-TSRI), Division of Plant Biology – BCC206, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA (*author for correspondence); [1]Present address: Université Paul Sabatier, UMR CNRS-UPS 5546, Centre de Biologie et Physiologie Végétale, 118 Route de Narbonne, 31062 Toulouse Cédex 4, France*

Received 12 December 1997; accepted in revised form 20 March 1998

Key words: cassava, caulimovirus, *cis* elements, GUS, promoter, transgenic

Abstract

Cassava vein mosaic virus (CsVMV) is a pararetrovirus that infects cassava plants in Brazil. A promoter fragment isolated from CsVMV, comprising nucleotides −443 to +72, was previously shown to direct strong constitutive gene expression in transgenic plants. Here we report the functional architecture of the CsVMV promoter fragment. A series of promoter deletion mutants were fused to the coding sequence of *uidA* reporter gene and the chimeric genes were introduced into transgenic tobacco plants. Promoter activity was monitored by histochemical and quantitative assays of β-glucuronidase activity (GUS). We found that the promoter fragment is made up of different regions that confer distinct tissue-specific expression of the gene. The region encompassing nucleotides −222 to −173 contains *cis* elements that control promoter expression in green tissues and root tips. Our results indicate that a consensus as1 element and a GATA motif located within this region are essential for promoter expression in those tissues. Expression from the CsVMV promoter in vascular elements is directed by the region encompassing nucleotides −178 to −63. Elements located between nucleotides −149 and −63 are also required to activate promoter expression in green tissues suggesting a combinatorial mode of regulation. Within the latter region, a 43 bp fragment extending from nucleotide −141 to −99 was shown to interact with a protein factor extracted from nuclei of tobacco seedlings. This fragment showed no sequence homology with other pararetrovirus promoters and hence may contain CsVMV-specific regulatory *cis* elements.

Introduction

Caulimoviruses (type member: cauliflower mosaic virus, CaMV) and badnaviruses (type member: commelina yellow mottle virus, ComYMV) are grouped together as plant pararetroviruses [34]. These viruses possess a circular double-stranded DNA genome of around 8 kb that is replicated through a process that includes a step of reverse transcription [30]. Replication is initiated by the synthesis of a terminally redundant genome-length RNA under the control of a viral promoter, also called the genomic promoter [34]. Pararetrovirus genomic promoters interact with the plant RNA polymerase II [27] and their regulation is apparently not dependent upon any virus-encoded transcriptional factors [26]. Consequently, these promoters were isolated from pararetrovirus genomes to express foreign genes in transgenic plants [6, 25]. The 35S promoter of CaMV has been widely used in plant biotechnology and its functional organization is well described. This promoter is highly active in plant cells and directs constitutive gene expression in most tissues of transgenic plants [15, 28, 40]. The regulatory region of the 35S promoter is made up of distinct domains that are each involved in organ-specific expression [1, 3]. It has been suggested that the strength and the constitutive properties of the 35S promoter de-

rive from additive and synergistic interactions between these different domains [2, 10]. Within these promoter domains two *cis*-acting elements were characterized. The activating sequence 1, or as1, is a TGACG repeat motif extending from −82 to −64 that is the binding site for the tobacco transcription factor TGA1a [16, 19]. It was shown that this *cis* element has an important role in the control of gene expression in roots [11, 19] and functions in synergy with upstream domains to potentiate promoter activities in other tissues [1, 3, 10, 11]. Also, a GATA motif located at from −85 to −105, called activating sequence 2 (as2), interacts with nuclear proteins from tobacco leaf tissues and has biological importance for promoter activity in those tissues [20]. A promoter similar to the 35S CaMV promoter was isolated from figwort mosaic caulimovirus, FMV [35]. This promoter, referred to as the 34S promoter, also directs constitutive gene expression in transgenic plants and is of comparable strength to the CaMV 35S promoter. Sequence comparison between the two promoters in the region extending from −100 to +1 reveals 65% nucleotide homology [35]. Based on the similarity between these two promoters, it has been suggested that similar mechanisms of regulation may govern their expression.

The cassava vein mosaic virus (CsVMV) is a double-stranded DNA virus that infects cassava in Brazil [17, 21] and has previously been listed as a putative member of the genus *Caulimovirus* [14]. However, the recent molecular characterization of the CsVMV genome [4, 7] revealed that this virus has a specific genome organization not found in other caulimoviruses or badnaviruses that have been characterized. Also it was proposed that CsVMV might be representative of a new genus of double-stranded DNA plant viruses [7]. Recently, we isolated a promoter fragment from the CsVMV genome [37] that is capable of driving high levels of heterologous gene expression in different plants and in protoplasts. This promoter directs expression of a reporter gene in all organs and in different cell types of transgenic plants suggesting a constitutive activity similar to that of the CaMV 35S promoter. However, the CsVMV has limited sequence homology with the two caulimovirus promoters previously described. These facts suggest the presence of different *cis*-elements in the promoter and therefore the possibility of different regulatory mechanisms compared with the CaMV 35S promoter. To address this question, we analyzed the functional organization of the CsVMV promoter using a deletion analysis of the region upstream of the TATA box. The effects of various deletions on promoter activity were monitored in transgenic tobacco plants using the *uidA* reporter gene. Additional data were also provided by transient expression experiments in protoplasts and by *in vitro* DNA-binding assays. We show that the constitutive pattern of expression of the CsVMV promoter is due to the presence of distinct organ specific elements. The relative importance of the different regions of the CsVMV promoter as well as its mechanisms of regulation are discussed.

Materials and methods

Promoter constructs and plasmids

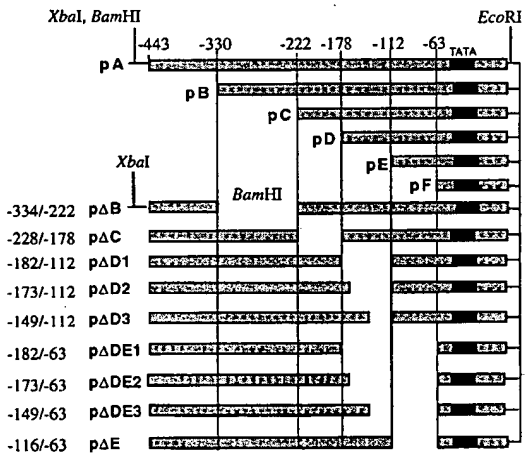
The starting plasmid for this study was pILTAB:CVP2 which contains a CsVMV promoter fragment extending from +72 to −443 [37]. Due to the absence of convenient restriction sites in the CsVMV promoter fragment, polymerase chain reaction (PCR) was used to generate a set of 5′-terminal and internal deletions.

Deletions from the 5′ end of the promoter were obtained by PCR amplification. A common reverse primer P1′ (Table 1) which hybridizes at the 3′ end of the promoter was paired with CsVMV-specific primers P2, P3, P4, P5 and P6 (Table 1) to generate five promoter fragments designated B, C, D, E and F. These fragments have a common 3′ terminus at position +72 and their 5′-end points at positions −330, −222, −178, −112 and −63, respectively (Figure 1). A full-length promoter fragment (A fragment) was also re-synthesized using the primer P1 and P1′ (Table 1). PCR reactions were carried out with 100 ng of pILTAB:CVP2, 2.5 U of *Taq* DNA polymerase (Gibco-BRL, Gaithersburg, MD) and standard concentrations of primers, $MgCl_2$ and dNTP. Twenty cycles (94 °C, 30 s; 56 °C, 30 s, 72 °C, 30 s) of amplification were performed and were followed by 5 min of final elongation at 72 °C. Each of the five amplified DNA fragments was digested by *Bam*HI and *Eco*RI and ligated into the same sites of a plasmid containing the coding sequence of the *uidA* gene (coding for β-glucuronidase) linked to the 3′ polyadenylation signal of the nopaline synthase gene. The resulting plasmids were named pA, pB, pC, pD, pE, pF according to the promoter deletion they carry (Figure 1).

The internal promoter deletions were constructed through a two-step process. First, PCR reactions were performed as described above to generate a set of 3′-end deletions of the CsVMV promoter. The P1 sense

Table 1. Oligonucleotides used to generate CsVMV promoter fragments by PCR amplification. Primers contain CsVMV promoter sequences in sense (+) or reverse orientation (−). Coordinates of the primers relative to the transcription start site are noted. P1′ in association with P2 to P6 were used to create 5′ terminal deletions of the CsVMV promoter as described in Materials and Methods. Similarly, P1 in association with P2′ to P7′ were used for 3′ end deletions. P1 and P1′ contain, respectively, a *Xba*I site and a *Eco*RI site at their 5′ ends while other primers encode *Bam*HI site. The restriction sites are indicated by bold letters. P7 and P8 as well as the reverse primers P8′ and P9′ were used to generate probes and competitor fragments for gel retardation assays (see Materials and methods for details).

Name	Primer sequence (5′ to 3′)	Position	Orientation
P1	GC**TCTAGA**CCAGAAGGTAATTATCCAAG	−443/ − 423	+
P2	TAT**GGATCC**TATGTTCAAAAATGAAG	−330/ − 312	+
P3	AAA**GGATCC**TGAAGACGTAAGCACTG	−222/ − 206	+
P4	AGA**GGATCC**GGTCGGTGATTGTGAA	−178 − /163	+
P5	AAA**GGATCC**TTATCACAAAGGAATC	−112/ − 95	+
P6	TAT**GGATCC**GTGTCATTTTTGCCCTTG	−63/ − 43	+
P7	GAGACATAGAGAGGACACATGT	−161/ − 142	+
P8	AAGGTGGAAAATGTAAGGGCG	−141/ − 121	+
P1′	CG**GAATTC**AAACTTACAAATTTCTCTAAG	+72/ + 50	−
P2′	TAA**GGATCC**TTTCCGCCCTTACATT	−116/ − 132	−
P3′	CAT**GGATCC**TCTATGTCTCTTTCAC	−149/ − 168	−
P4′	ACA**GGATCC**GACCTTATCTTCT	−173/ − 187	−
P5′	ACC**GGATCC**TCTTCTTTTCATTGTTC	−182/ − 199	−
P6′	TCA**GGATCC**TTTTCTTCGCCTGGT	−228/ − 243	−
P7′	ATA**GGATCC**ATATGTGCCGCATA	−334/ − 348	−
P8′	CCTTTGTGATAAGGTTACTTTCC	−99/ − 121	−
P9′	GACACGGAAAAATATAAAAGG	−56/ − 76	−



	Mesophyll	Vascular	Root tips
pA	+	+	+
pB	+	+	+
pC	+	+	+
pD	(-)	+	-
pE	-	(+/-)	-
pF	-	-	-
-334/-222 pΔB	+	+	+
-228/-178 pΔC	(+/-)	+	+
-182/-112 pΔD1	-	(+/-)	+
-173/-112 pΔD2	+	+	+
-149/-112 pΔD3	+	+	+
-182/-63 pΔDE1	-	(-)	+
-173/-63 pΔDE2	-	(+/-)	+
-149/-63 pΔDE3	-	(+/-)	+
-116/-63 pΔE	+	+	+

Figure 1. Schematic representation of the deleted CsVMV promoter:*uidA* fusion genes and a summary of histochemical detection of GUS activity in transgenic tobacco plants. The names of the different plasmids and the end points of the 5′ and internal deletions are indicated. Internal deletions are designated by symbol Δ. The construct pA contains a full-length CsVMV promoter. All promoters deleted from the 5′ end have a *Bam*HI site at their 5′ ends. Internal deletions were created by *Bam*HI ligation of the promoters truncated from the 5′ end with those truncated from the 3′ end (see Materials and methods). The data summarize the observations performed on ten different transgenic R1 tobacco plants for each promoter construct. The transgenic plants were analyzed for GUS expression in mesophyll cells, vascular tissues and root tips. GUS expression is scaled in each tissue according to 4 levels. +: no visible difference with the full length promoter; (+/−): lower staining than with the full length promoter; (−): very low expression; −: no detectable staining.

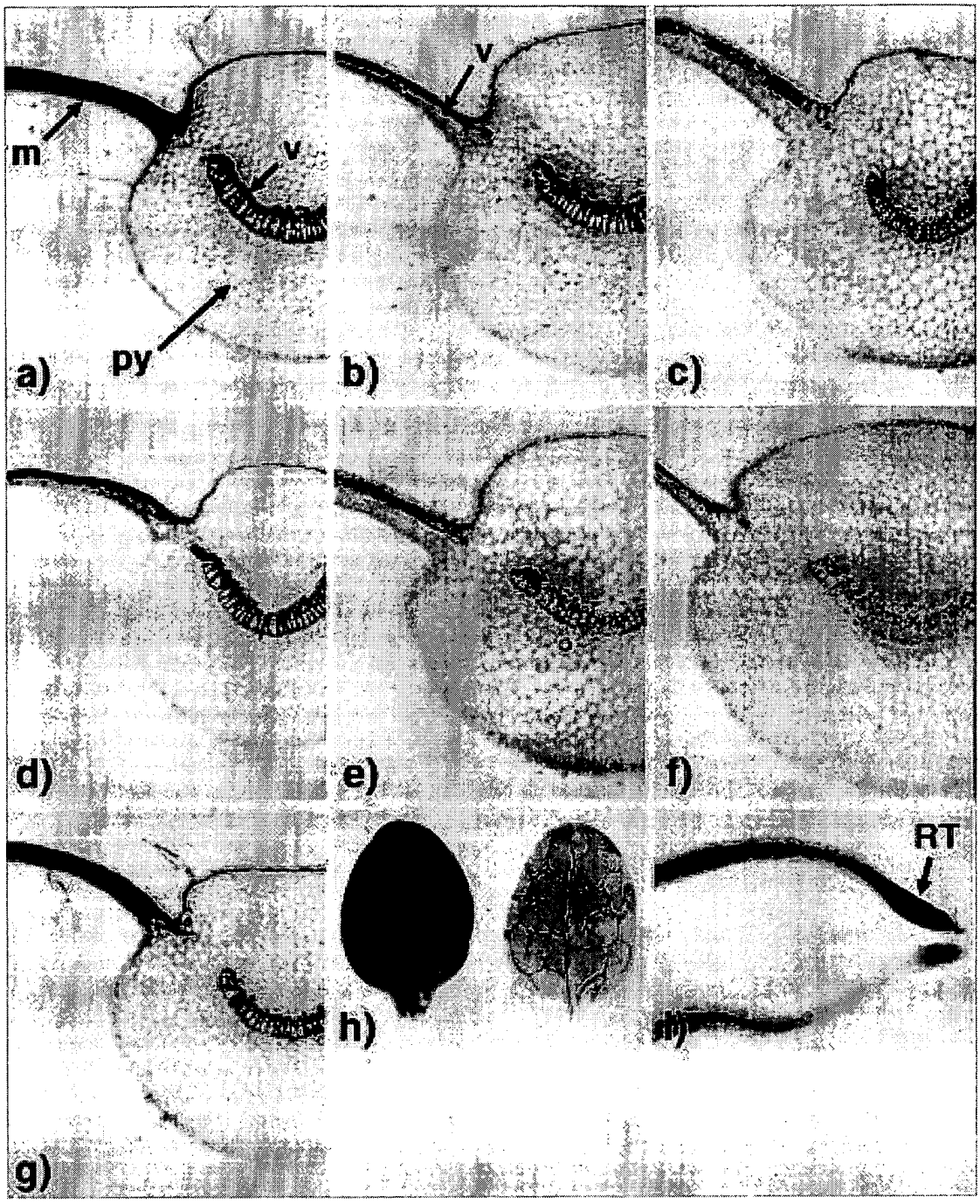

m
v
py
a)
v
b)
c)
d)
e)
f)
RT
g)
h)

Figure 2. Histochemical localization of GUS expression in transgenic tobacco plants containing deleted CsVMV promoters:*uidA* fusion gene constructs. Pictures presented, with the exception of h and i, are cross-sections of young expanded leaves from 5-week old transgenic R1 tobacco plants. a. Construct pB. This expression profile in leaf was also observed with plants containing the construct pC. As with the full-length CsVMV promoter, strong GUS expression was detected in photosynthetic tissues (palisade and spongy mesophyll) as well as in the phloem cells and in the xylem parenchyma of the midrib. b. Construct pD. Strong staining was localized in the lateral secondary veins, in the phloem tissue and in the xylem parenchyma of the midrib while no or weak expression was detected in the mesophyll tissues. c. Construct pE. Secondary lateral veins as well as the vascular elements of the midrib exhibited a weak staining. d. Construct pΔC. Strong staining was detected in the vascular tissues of the midrib and in the lateral veins. We noticed a decrease in GUS expression in photosynthetic tissues as shown on the picture by a moderate staining in spongy mesophyll cells. e. Construct pΔD1. The same GUS expression pattern was observed with the constructs pΔDE2 and pΔDE3. Moderate staining is visualized in the lateral veins and in the phloem cells of the midrib. f. Construct pΔDE1. No expression was detectable with the exceptions of few phloem cells that exhibited a very weak staining. g. Construct pΔD2. Plants carrying the deletion constructs pΔD3 and pΔE displayed similar GUS expression pattern. GUS expression was detected in the mesophyll and vascular tissues and was in the same range of intensity as what we observed with the undeleted promoter. h. Leaves from 10-day old transgenic seedlings carrying pB construct (left) or pD construct (right). i. Roots from transgenic tobacco plants containing the pC promoter construct (top) or pD promoter construct (bottom). m, mesophyll; py, parenchyma, v, vascular tissue.

primer (Table 1) which hybridizes at the 5′ end of the promoter was paired with each of 6 specific CsVMV reverse primers, P2′ to P7′ (Table 1), to generate six truncated promoters with a common 5′ end at position −443 and 3′ end points spanning from position −116 to −334. Internal deletions were engineered by cloning the different 3′ end truncated promoter fragments upstream of the 5′ end deleted promoter constructs (Figure 1). Accordingly, a 3′ deleted promoter fragment encompassing nucleotides −443 to −334 was digested by *Bam*HI and *Xba*I and ligated to the same sites in the pC plasmid. The resulting plasmid, named pΔB (Figure 1) contains an internal deletion from nucleotides −334 to −222 (Figure 1). Similarly, a fragment spanning nucleotides −443 to −228 was fused to the D promoter fragment to create the plasmid pΔC (Figure 1). Three fragments with a common 5′ end at −443, and 3′ ends located at −182, −173, and −149 were cloned individually into the plasmid pE to create the plasmid pΔD1, pΔD2 and pΔD3, respectively (Figure 1). The same three fragments were cloned into pF to create the plasmid pΔDE1, pΔDE2 and pΔDE3 (Figure 1). A fragment containing nucleotides −443 to −116 was cloned with the same method into the plasmid pF to generate the plasmid pΔE (Figure 1). All promoter sequences were verified by dideoxynucleotide sequencing. The different CsVMV promoter-*uidA* fusion genes were excised by *Xba*I and *Hin*dIII and ligated to the same sites of the pBin 19 binary vector (Clontech, Palo Alto, CA).

Transgenic tobacco plants

The pBin 19 derivatives carrying the deleted promoter constructs were transferred by electroporation into *Agrobacterium tumefaciens* strain LBA4404. *Agrobacterium*-mediated transformations of *Nicotiana tabacum* cv. Xanthi NN were performed as previously described [13]. Regenerated kanamycin resistant plants were grown to maturity in a greenhouse and allowed to self-fertilize. R1 seeds were germinated on Murashige and Skoog (MS) culture medium [24] with 100 mg/l kanamycin and transferred to soil in a greenhouse. Between 10 and 20 independent transgenic lines were produced for each construct. Ten independent R1 lines for each promoter construct were analyzed for this study.

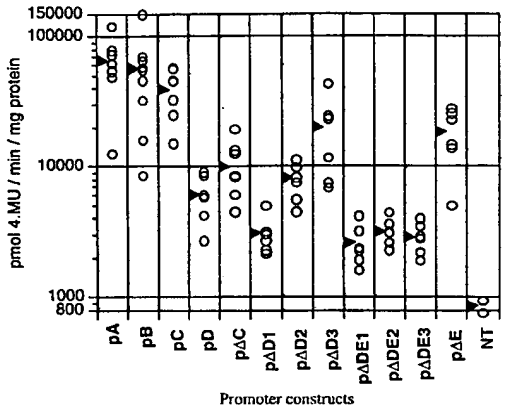


Figure 3. GUS enzyme activity in transgenic tobacco leaves carrying CsVMV promoter constructs. Proteins were extracted from leaf disks collected from young expanded leaves of 5 week old transgenic tobacco plants. For each promoter construct, 6 to 10 independent transgenic tobacco lines were assayed for GUS activity. Data are given as pmol of 4-methylumbelliferone per mg of protein per minute. The results obtained for each plant is represented by an open dot. The different promoter constructs analyzed in these assays are indicated. Average level of expression is indicated by a vertical arrowhead. A logarithmic scale is used to accommodate the large variation between transgenic lines.

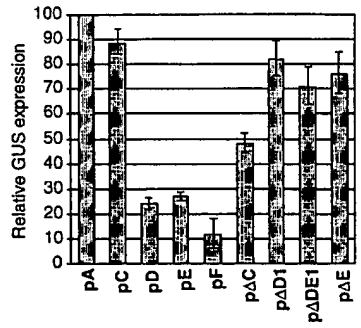


Figure 4. Transient GUS expression of CsVMV:*uidA* gene constructs in BY-2 protoplasts. Electroporated protoplasts prepared from BY-2 cell suspensions were analyzed for GUS activity after 24 h of culture. The deleted CsVMV promoter constructs were co-transfected with a luciferase plasmid used as an internal standard. GUS expressions levels were normalized in relation to the activity of luciferase. Normalized values of GUS expression are expressed as a percentage of the full-length promoter activity. A 100% value was assigned to the construct pA which contains an undeleted CsVMV promoter. The different promoter constructs are indicated on the figure. Bars represent the average of four independent experiments +/− standard errors.

GUS expression analysis

For the histochemical detection of β-glucuronidase (GUS) activities, fresh tissue sections were taken from young expanded leaves and incubated for 6 to 12 h at 37 °C in reaction buffer [15] containing 1 mM 5-bromo-4-chloro-3-indolylglucuronide (X-gluc), 100 mM sodium phosphate buffer pH 7, 2 mM potassium ferrocyanide and potassium ferricyanide, and 0.1% Triton X-100. For the GUS analysis of young R1 seedlings, whole plantlets were collected around one week after germination and immersed in GUS buffer [15]. After vacuum infiltration, incubation was carried out overnight at 37 °C. Samples were cleared by several washes in 70% ethanol. Quantitative GUS analyses using the substrate 4 methylumebelliferone-β-S glucuronide (MUG) were performed as described by Jefferson *et al.* [15].

Protoplast isolation, treatment and culture

Protoplasts from BY-2 tobacco suspension cells were prepared essentially as described by Watanabe *et al.* [39]. A 30 μg portion of plasmid DNA with 30 μg of a 35S-luciferase construct [28], which served as internal standard, was introduced into protoplasts as previously described [37]. Protoplast were collected for protein extraction after 24 h of culture at 25 °C. The MUG and LUC assays were performed on the protoplast protein extracts as described [37]. Results were expressed as a ratio between the GUS activity of the CsVMV promoter construct and the LUC activity of the internal control.

Probes and competitors

DNA fragments used as probes in binding assays were generated by PCR using 200 ng of pILTAB:CVP2 plasmid containing the full-length CsVMV promoter [37], *Pfu* DNA polymerase (Stratagene, La Jolla, CA) and its appropriate buffer, specific primers, and dNTPs as standard concentrations. A DNA fragment encompassing −161 to −56 was obtained using the primers P7 and P9′ (Table 1) after 20 cycles of amplification performed with the following temperature profile: 30 s at 94 °C; 30 s at 58 °C; 45 s at 72 °C. The PCR mixture was extracted with phenol/chloroform and the DNA was ethanol-precipitated. Amplified DNA was subjected to electrophoresis through on a 3% Biogel (Bio 101, La Jolla, CA) and the probe fragment was eluted using the Mermaid kit (Bio 101). Purified DNA was end-labeled with [γ^{32}P] ATP using T4 polynucleotide kinase (NEB, Beverly, MA). The probe was purified from unincorporated nucleotides by two successive rounds of ethanol precipitation in the presence of ammonium acetate.

Specific DNA competitors were produced by PCR using *Pfu* polymerase and the conditions described above. In addition to the −161/−56 fragment, three subsets encompassing −161 to −99 (obtained with P7 and P8′ primers, Table 1), −141 to −56 (obtained with P8 and P9′ primers) and −141 to −99 (obtained with P8 and P8′ primers) were used as competitors. DNA fragments were gel purified using the Mermaid kit (Bio 101). A 557 bp DNA fragment isolated by *Eco*RI and *Eco*RV digestion of the coding region of the *uidA* gene was used as a non-specific competitor DNA. This fragment contains *uidA* nucleotides from the ATG position to the first *Eco*RV site in the sequence.

Preparation of tobacco nuclear extracts

A protein nuclear extract was prepared from 2-week old tobacco (*N. tabacum* cv. Xanthi NN) seedlings grown in growth chambers (16 h day light cycle). All purification steps were carried out at 4 °C. Nuclei were isolated following the protocol described by Giuliano *et al.* [12] except that benzamidine was omitted from the extraction buffer. The isolated nuclei were then resuspended (50 μl/g of starting tissue) in lysis buffer

containing 20 mM HEPES pH 8, 0.4 mM EDTA, 0.5 mM DTT, 420 mM NaCl, 25% (v/v) glycerol, 0.1 mM PMSF, 2 μg/l leupeptine, 3 μg/l pepstatine and 40 μg/l antipain. The suspension of nuclei was homogenized by pipeting up and down and incubated on ice for 1 h (with occasional mixing). The resulting extract was centrifuged at 100 000 × *g* for 1 h. The supernatant was collected and dialyzed for 2 h against 2 liters of dialysis buffer containing 20 mM HEPES pH 8, 50 mM KCl, 0.4 mM EDTA, 20% glycerol, 0.1 mM PMSF and 0.5 mM DTT. The dialysate was centrifuged at 12 000 × *g* for 10 min to remove insoluble material and the supernatant was aliquoted and stored at −80 °C. Protein concentrations of nuclear extracts were determined using the Bradford assay (BioRad, Hercules, CA).

Gel shift assays

Protein-DNA binding reactions were carried out on ice in 25 μl of 1 × binding buffer (20 mM HEPES pH 7.8, 50 mM KCl, 1 mM $MgCl_2$, 0.05% Nonidet 40, 1 mM DTT and 10% glycerol) containing 2 μg of poly-dI-dC as non-specific competitor and an adequate amount of nuclear extract protein (amounts of nuclear extract are noted in the legend to Figure 5). The mixture was pre-incubated for 10 min on ice before addition of 10 000 cpm of the labeled probe. The reaction was allowed to proceed on ice for an additional 25 min.

For competition assays, competitor DNAs were included in the reaction before addition of the labeled probe. The amount of competitor used for each assay is indicated in Figure 5. Reaction mixtures were subjected to electrophoresis through a 5% non-denaturing polyacrylamide gel in 1× TGE buffer, pH 8.5, at 130 V for 3 h at 4 °C. Gels were fixed in 10% methanol, 10% acetic acid, dried and exposed overnight on X-ray films with an intensifying screen.

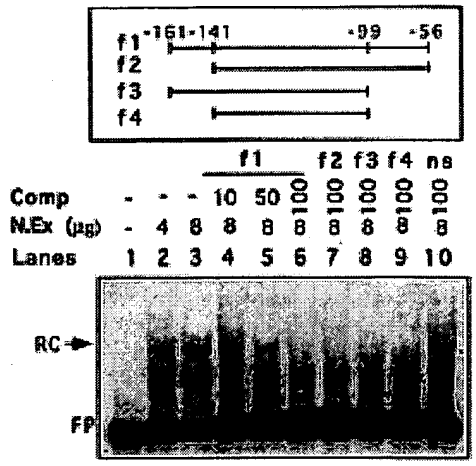


Figure 5. Interaction of a fragment extending from nucleotide −161 to −56 with nuclear extracts from tobacco seedlings. The fragment f1, spanning nucleotides −161 to −56, was end-labeled and incubated in the presence or absence (−) of nuclear extract (N.ex) as indicated in Materials and methods. The free probe (FP) is visible at the bottom of the figure while a DNA/protein complex (RC) of slower electrophoretic mobility is marked by an arrow. In lanes 4 to 10, 10 to 100× molar excess of unlabeled competitor DNA were added to the binding reaction mixtures. In addition to the f1 fragment, f2, f3 and f4 fragments were used as competitors. In lane 10, 100× molar excess of a 557 bp fragment isolated from the *uidA* coding sequence was used as a non-specific competitor.

Results

Expression patterns of deleted CsVMV promoter constructs in transgenic tobacco plants

The CsVMV promoter was mutated by progressive 5′ deletions and by internal deletions as described in Materials and methods. Expression patterns of the different promoter constructs were analyzed in transgenic R1 tobacco plants using histochemical staining of GUS activity. The presence of an intact promoter:*uidA* gene cassette was confirmed by PCR and/or Southern analysis for each transgenic plant line used in this study (data not shown). For each promoter construct, we obtained between 10 to 20 transgenic plant lines. GUS expression patterns observed between plants containing the same promoter construct were similar while we observed variations in the level of transgene expression among individual transformants. The plants tested in this study contained between 1 and 5 copies of the *uidA* fusion gene (data not shown). The copy number did not affect the characteristic pattern of expression observed with each construct and there was no clear correlation between the copy number and the apparent intensity of the staining. For each promoter construct, we always obtained a small number of transgenic plants for which the expression of the transgene was not detected despite its presence as shown by Southern blotting, suggesting a possible silencing phenomenon [29]. Data obtained from these plants were not taken into account for the current deletion analysis. However, in the lines selected for this study, transgene expression was stable during plant development and the pattern of GUS staining was similar

in the primary transformants and in the R1 generation indicating that silencing phenomena did not occur in the selected plant lines [9].

Different patterns of expression as well as significant and reproducible differences in the staining intensities between promoter constructs could be clearly visualized despite some variation between individual transformants. It is likely that the reproducible differences detected between promoter constructs indicated an effect of the deletion on the promoter regulation. We reported previously that the CsVMV promoter is expressed in all organs of transgenic plants, with regions of highest promoter expression located in vascular tissues, mesophyll cells of the leaves and the root tips [37]. Accordingly, for the present study, GUS activity was analyzed in these three different tissues (Figure 1).

5′-end deletions

GUS staining in transgenic plants that contained the promoter constructs deleted to −330 and −222 (constructs pB and pC) occurred in the same pattern and was in the same range of intensity as what was observed with the full length promoter (Figures 1 and 2a). Further deletion of the promoter to −178 (construct pD) caused an important change in the GUS expression pattern (Figure 1). In leaf cross-sections of most of the plants carrying the pD construct, a strong staining restricted to the vascular tissues was observed (Figure 2). GUS staining was absent in the mesophyll cells of most of the plants. However, three plants lines out of ten presented a low staining in these cells. In all plants transformed with the pD construct, the root tips did not exhibit GUS staining (Figure 2i) while this tissue was intensely stained in plants containing the pC promoter construct. GUS activity from the promoter construct pE, deleted to −112, was restricted to the vascular tissues (Figures 1 and 2c). The intensity of the expression was very low and long incubation times were required to detect a blue precipitate in these tissues. The construct pF did not display any detectable expression (Figure 1).

Internal deletions

Internal deletion of the region from −334 to −222 (promoter ΔB) did not affect the general expression pattern of the CsVMV promoter (Figure 1). A significant decrease in GUS accumulation was visualized in the spongy and palisade mesophyll cells of the plants transformed with pΔC (Figures 1 and 2d). However the deletion did not suppress expression in vascular elements and root tips, which exhibited a strong staining. Transgenic plants containing pΔD1 (deletion from −182 to −112) displayed a weak GUS staining restricted to the phloem cells (Figures 1 and 2e). Additionally, GUS expression was also observed in the root tips. The internal deletion encompassing −182 to −63 (pΔDE1) had a profound effect on promoter activity (Figure 1). Of the 10 independent transgenic plants tested, 8 did not have any detectable GUS activity in the leaves or in stems. Very pale blue staining localized in the phloem cells was visualized in two plants after prolonged incubation (Figure 2f). In contrast, strong staining was revealed in the root tip as well as weaker staining in the vascular elements of the root. The construct pΔE displayed a constitutive pattern of expression similar to the undeleted promoter (Figures 1 and 2g).

These results emphasized the importance of the (−182/−112) region for constitutive expression of the CsVMV promoter. We observed that the deletion pΔD1 suppressed promoter activity in the mesophyll tissue and also diminished vascular GUS expression (Figure 2e). Consequently, we made the constructs pΔD2 and pΔD3 to investigate this region in greater details (Figure 1). The construct pΔD2, deleted from nucleotides −173 to −112, displayed a profile of expression similar to the full length CsVMV promoter (Figure 2g). This result suggested that the addition of 9 nucleotides at the 5′ end point of the (−182/−112) deletion could restore the full expression pattern altered with the construct ΔD1. No significant difference was detected between the construct pΔD2 and pΔD3.

In contrast, addition of the sequences −182/−173 and −182/−149 to pΔDE1 to make pΔDE2 and pΔDE3, respectively, did not lead to a restoration of promoter activity (Figures 1 and 2e), suggesting that elements located between −112 and −63 might influence promoter activity.

Histochemical GUS analysis in young seedlings

A histochemical GUS analysis on 10-day old R1 seedlings was carried out in order to analyze the expression pattern of the deleted promoter at earlier stages of development. The deleted promoter constructs directed specific expression patterns that were similar to those observed in adult plants. As in expanded leaves of adult plants, the constructs pD, pE and pΔD1 conferred a vascular-specific pattern of expression in cotyledonary leaves (Figure 2h). Sim-

ilarly, the pB and pC constructs which were active in both mesophyll and vascular tissues in tobacco plants had the same constitutive expression pattern in seedlings (Figure 2h). These results suggested that specific expression patterns observed with the different deletions in transgenic plants are not influenced by the developmental stage of the plant.

GUS activity in leaf extracts of transgenic plants

GUS activities in protein extracts prepared from leaf tissues were measured using a fluorometric assay [15]. The samples were collected from interveinal tissue of young expanded leaves from 5-week old transgenic tobacco plants. Consequently, the GUS activity levels reflected mainly the promoter expression in mesophyll tissue. As shown in Figure 3, the values of GUS activities in transgenic lines carrying a specific promoter construct varied by a maximal factor of 17. We did not observe a clear correlation between levels of expression and transgene copy number (not shown) as plants with the same copy number exhibited large differences in levels of GUS activity.

A significant decrease in level of GUS expression was detected in transgenic plants containing the pD promoter construct. A low level of GUS activity was also measured with the internal deletion pΔC confirming the importance of the region extending from −222 to −178. This result was consistent with the histochemical analysis since pD and pΔC displayed low or no staining in mesophyll cells of transgenic plants. Similarly, low levels of GUS activity were measured for the constructs pΔD1, pΔDE1, pΔDE2 and pΔDE3 which had vascular-specific profiles of expression in leaves of transgenic plants. The average level of activity measured with the construct pΔD2 was higher than with the pΔD1 construct confirming that the region from −182 to −173 is involved in promoter activation. However, the pΔD3 activity levels were in the higher range while the construct pΔD2 gave moderate levels of expression. This difference was not detected using the histochemical assay (Figure 2g). As observed with the histochemical assay, the high expression level measured with the construct pΔD3 was abolished by deletion of the region spanning −112 to −63 (construct pΔDE3). The construct pΔE retained a high level of GUS activity suggesting that the promoter expression in green tissues is weakly affected by this deletion.

Expression of the CsVMV promoter deletions in protoplasts

Protoplasts prepared from BY-2 suspension cells were transfected with the CsVMV promoter constructs. Transient expression of the *uidA* reporter gene was measured 24 h after electroporation in relation to an internal standard expressed from a cotransfected luciferase plasmid. Four independent transfection experiments were carried out; the results obtained are summarized in Figure 4.

The construct pC which contains a CsVMV promoter deleted to position −222 retained 88% of the activity of the full-length promoter fragment. The promoter activity dropped sharply to only 24% of full activity with a further deletion to −178. Constructs pD and pE had almost the same expression level while a second drop of activity was observed with a deletion extending to −63. The construct pF with 12% of the full promoter activity was just above background level. An internal deletion from −228 to −178 (construct pΔC) decreased expression by more than 50%. Surprisingly the constructs pΔD1 and pΔDE1 which gave very low expression in transgenic plants allowed high levels of expression in protoplasts. The (−222/−178) region plays a critical role for promoter activity in protoplast while the region from −178 to −63 appears to be of lower importance for expression of the promoter in protoplasts than in plants.

DNA-protein interactions in the −161 to −56 region

A labeled DNA fragment encompassing −161 to −56 was used to detect specific binding with nuclear protein extracts prepared from tobacco seedlings. Interaction of the (−161/−56) probe with the nuclear extract generated a single complex of slower mobility than unbound probe (Figure 5). The intensity of the retarded band increased with increasing amount of nuclear extract in the reaction. The formation of the complex was reduced or abolished by addition of an excess of unlabeled probe whereas addition in the reaction of a an unrelated DNA fragment had no effect (Figure 5). This result suggested that one or more DNA-binding proteins in the tobacco nuclear extract interact with *cis*-elements within the (−161/−56) promoter region.

To map more precisely the binding site of this interaction, we performed competition assays with sub-fragments of the probe. The complex was eliminated when overlapping sub fragments spanning −161 to −99, −141 to −56 or −141 to −99 were added to the reactions at 100-fold molar excess in the assay.

This suggested that a binding site for a protein factor is located between −141 and −99.

Discussion

This study showed that the CsVMV promoter has a complex structure made of discrete domains that exercise distinct influences on patterns of tissue specific expression. Our results suggest that the constitutive activity of the promoter is the result of additive and synergistic interactions between different promoter elements. The functional organization of the CsVMV promoter as defined by the deletion analysis is summarized in Figure 6.

Promoter expression in green tissues is directed by the region extending from −222 to −173. In this region, we identified a consensus sequence as1 [19] at −219/−203 as well as a GATA sequence at position −183/−180. The construct pΔD1 (deletion of the region from −182 to −112) exhibited a low level of expression (Figure 2e) and showed that as as1 element on its own cannot activate gene expression in mesophyll tissue. As suggested when we analyzed the expression patterns of the construct pΔD1 and pΔD2, the sequence extending from −182 to −173, which contains a GATA motif, is essential for promoter activity in green tissue. However, the specific role of this GATA region independent of the as1 element cannot be assessed based on expression of the constructs used in this study. Consequently, two hypotheses are presented: either the GATA region, on its own, controls promoter expression in green tissues, or the GATA region and the as1 element act together to control the CsVMV promoter activity in these tissues. Published data reported that a GATA motif in the CaMV promoter, referred to as the activating sequence 2, as2 [20], is also involved in leaf expression. Moreover, gene expression in leaf tissues controlled by the GATA region in the CaMV promoter was dependent on sequences located within the −90 to −46 region (which contains as as1 element). Similar interactions may control expression of the CsVMV promoter in green tissues. However, the GATA motif identified in the CsVMV promoter is not identical with the as2 element of the CaMV promoter. We found a greater sequence homology with a GATA box identified in the rice tungro bacilliform badnavirus promoter [5, 42] which plays an important role in the activation of that promoter [43] and with a box I element (GATAAPu) found in several light and circadian-clock-regulated promoters [8, 36].

A low level of GUS expression was observed in mesophyll cells of plants transformed with the promoter construct pΔC which deleted the as1 and GATA regions (Figure 2d). It is unlikely that this residual GUS staining in green tissue is due to additional minor elements located in the region from −443 to −228. Indeed, this region in the context of the deletion pΔD1 did not direct any GUS expression in mesophyll cells (Figure 2e). The GUS staining in mesophyll cells of plants containing pΔC is due either to diffusion of the dye from intensely colored vascular tissues or to the new arrangement of sequence created by the deletion that could be favorable to gene expression in the mesophyll cells.

The results of the studies in transgenic plants showed that expression of the CsVMV promoter in vascular tissues is directed by the region encompassing −178 to −63 and is not dependent upon the presence of as 1 or the GATA region (as shown by the construct pD, Figure 2b). The latter elements may have a moderate effect on vascular expression as suggested by the construct pΔDE2 and pΔDE3 (Figure 2e) but is clearly not required for full promoter activity in vascular tissues. The expression pattern displayed by construct pE showed that a vascular domain is located in the −112/−63 region. We detected in this region a 22 nucleotide sequence containing a CTTATC repeat previously identified in the ComYMV promoter [23]. However, the results obtained showed that additional elements located between −178 to −112 are responsible for the strong vascular expression (Figure 6). Within the latter region, important elements may be contained between −149 to −116 as suggested by comparing the expression pattern of pΔDE3 and pΔE.

In vitro binding assays performed with the region −161 to −56 revealed a specific interaction with nuclear protein. We detected a single retarded band in these assays, the formation of which was efficiently disrupted by competition with a 43 nucleotide fragment extending from −141 to −99. This was somewhat surprising since regions from −149 to −112 and from −112 to −63 play an active role in expression of the promoter. It is possible that certain binding activities are not detected due to a low concentration of transcriptional factors in our nuclear extract, to a lower affinity for the binding site, or because cooperative binding with other factors is required.

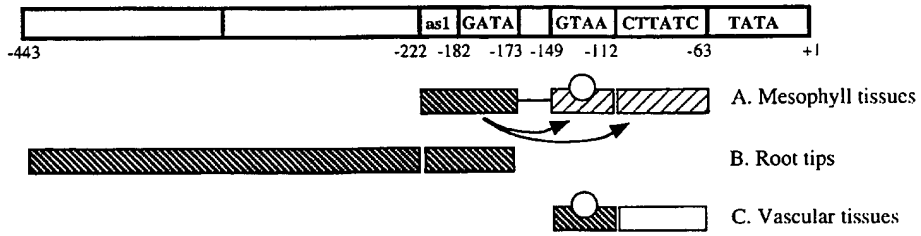


Figure 6. Summary of the CsVMV promoter regions required for *uidA* gene expression in various tissues of transgenic tobacco plants. Coordinates relative to the transcription start site indicate the position of promoter domains as defined in this study. Promoter domains required for expression in mesophyll tissues (A), in root tips (B) and in vascular tissue (C) were identified and are indicated below the map. A. An essential promoter region for expression in mesophyll tissues (hatched grey box) is located between −222 and −173. This region contains a consensus asl element [19] and a GATA motif. The hatched boxes from −149 to −112 and from −112 to −63 represent promoter regions that could not specify mesophyll expression by themselves but that were required in combination with the upstream hatched grey box to direct expression in green tissues. These putative synergistic interactions are represented in the figure by the arrows. B. Two promoter domains (hatched grey boxes) from −443 to −222 and from −222 to −178 could direct gene expression in root tips. Our data suggest that these two domains have a redundant function. C. The region extending from −149 to −112 (hatched grey box), characterized by a GTAA repeat, was responsible for high level of gene expression in vascular tissues. The region extending from −112 to −63 (open box) was also able to specify vascular expression but could not direct high level of gene expression. This region contains a CTTATC repeat previously identified within the ComYMV promoter sequence [23]. Formation of a complex between the promoter region extending from −141 to −99 and a protein factor present in tobacco nuclear extract is indicated by an open circle.

More detailed examination of the sequence of this 43 nt fragment revealed the presence of a GTAA repeat located at −129 to −113. The upstream GTAA motif of the latter repeat resembles the core-sequence of the P-box identified in the zein gene promoters [38]. The *tef*1 box [32] located in the −100 region of the EF-1a gene promoters contains a GTAA repeat with similarities to the GTAA box of the CsVMV promoter. Determination of the role of the GTAA repeat in the CsVMV promoter will require further study.

Deletion of regions of the CsVMV promoter that are important for expression in vascular tissues, as in the construct pΔDE3, suppressed gene expression in green tissues. Expression in the mesophyll tissue was restored by the addition of nucleotides −419 to −116 (construct pΔE) or by the addition of nucleotides −112 to −63 (construct pΔD3). These results could be explained by an effect due to spacing between the 'mesophyll domain' (namely, the asl element and the GATA motif) and the TATA box in construct pΔDE3. By contrast, the promoter is active when at least 49 nucleotides are present between the GATA element (−182 to −173) and nucleotide −63.

In the CaMV 35S promoter, the asl and GATA motifs are located between −64 and −105 [20], much closer to the TATA box than in the CsVMV promoter. By analogy, the shorter distance between the GATA region and TATA box created by the pΔDE2 and pΔDE3 internal deletions would not be expected to prevent the activity of the asl and GATA *cis* elements. The data may suggest that additional *cis*-acting elements contained in the region −149/−112 and in the region −112/−63 are required for promoter activity in green tissues (Figure 6). Elements contained in these regions may act in synergy with the 'mesophyll domain' of the CsVMV promoter to promote gene expression in green tissues.

The CsVMV promoter includes an important regulatory domain for expression in root tips between −222 and −178 (Figure 2i). It is likely that the asl element contained in this region is involved in this function, based on analogy with the role of the asl element of the 35S CaMV promoter in root expression [1, 11, 19]. However the construct pΔC, which lacks the asl element, is strongly expressed in root tip, and indicates that other promoter-elements are also involved in root tip expression (Figure 6).

In BY2 protoplasts, CsVMV promoter activity appeared to be controlled by the region encompassing −222 to −178 that contains the asl consensus sequence. Similar results were obtained with leaf mesophyll protoplasts (not shown). It was surprising to find that expression from the promoter is independent of sequences from −182 to −63. Discrepancies of results between protoplast-based transient assays and transgenic plants were also observed for the 35S CaMV promoter [10, 18]. Protoplasts are in highly stressed physiological state [33] that could be responsible of activation or inactivation of various *trans*-acting factors interacting with the promoter. In this regard,

several reports on the responsiveness of the as1 element to multiple stress-related signals such as auxins, salicylic acid, methyl jasmonate are particularly relevant [22, 31, 41, 44]. Interestingly, the activities of the promoter mutants in protoplasts correlates with their activities in root tips (Figures 1 and 4) suggesting that a similar array of *trans*-acting factors is present in root tips and in protoplasts.

Acknowledgements

We thank Sally Jo L. Leitner for great assistance in plant maintenance. We are grateful to Dr Hal S. Padgett for critical reading of the manuscript. This work was supported by the Rockfeller Foundation and by the French Institute of Scientific Research for Development in Cooperation (ORSTOM).

References

1. Benfey PN, Chua NH: CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue specific expression patterns. EMBO J 8: 2195–2202 (1989).
2. Benfey PN, Chua NH: The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants. Science 250: 959–966 (1990).
3. Benfey PN, Ren L, Chua NH: Combinatorial and synergistic properties of CaMV 35S-enhancer subdomains. EMBO J 9: 1685–1696 (1990).
4. Calvert L, Ospina M, Shepherd R: Characterization of cassava vein mosaic virus: a distinct plant pararetrovirus. J Gen Virol 76: 1271–1276 (1995).
5. Chen G, Muller M, Potrykus I, Hohn T, Futterer J: Rice tungro bacilliform virus: transcription and translation in protoplasts. Virology 204: 91–100 (1994).
6. Covey S, Hull R: Genetic engineering with double-stranded DNA viruses. In: Wilson TMA, Davies JW (eds) Genetic Engineering with Plant Viruses, pp. 217–221. CRC Press, Boca Raton, FL (1992).
7. de Kochko A, Verdaguer B, Taylor N, Carcamo R, Beachy RN, Fauquet C: Cassava vein mosaic virus (CsVMV), the type species for a new genus of plant pararetroviruses. Arch Virol, in press.
8. Donald G, Cashmore A: Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis* rbcs-1A promoter. EMBO J 9: 1717–1726 (1990).
9. Elmayan T, Vaucheret H: Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionaly. Plant J 9: 787–797 (1996).
10. Fang RX, Nagy F, Sivasubramaniam S, Chua NH: Multiple *cis*-regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell 1: 141–50 (1989).
11. Fromm H, Katagiri F, Chua NH: An octopine synthase enhancer element directs tissue-specific expression and binds ASF-1, a factor from tobacco nuclear extracts. Plant Cell 1: 977–984 (1989).
12. Giuliano G, Picherski E, Malik VS, Timko MP, Scolnik PA, Cashmore AR: An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci USA 85: 7089–7093 (1988).
13. Horsch RB, Fry J, Hoffmann N, Neidermeyer J, Rogers SG, Fraley RT: Leaf disc transformation. In: Gelvin SB, Schilperoort RA (eds) Plant Molecular Biology Manual, pp. A5/1–A5/9. Kluwer Academic Publishers, Dordrecht, Netherlands (1988).
14. Hull R: Caulimovirus. In: Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds) Virus Taxonomy. Sixth report of their international committee on taxonomy of viruses, pp. 189–192. Springer-Verlag, Wien/New York (1993).
15. Jefferson RA, Kavanagh TA, Bevan MW: GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901–3907 (1987).
16. Katagiri F, Lam E, Chua NH: Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. Nature 340: 727–730 (1989).
17. Kitajima EW, Costa AS: Particulas esferoidais associadas ao virus do mosaico das nervuras da mandioca. Bragantia 25: 211–221 (1966).
18. Lam E: Analysis of tissue-specific elements in the CaMV 35S promoter. Results Probl Cell Differ 20: 181–196 (1994).
19. Lam E, Benfey PN, Gilmartin PM, Fang RX, Chua NH: Site-specific mutations alter *in vitro* factor binding and change promoter expression pattern in transgenic plants. Proc Natl Acad Sci USA 86: 7890–7894 (1989).
20. Lam E, Chua NH: ASF-2: a factor that binds to the cauliflower mosaic virus 35S promoter and a conserved GATA motif in Cab promoters. Plant Cell 1: 1147–1156 (1989).
21. Lin MT, Kitajima EW: Purificao e serologia do virus do masaico das nervuras da mandioca. Fitopatol Brasil 5: 419 (1980).
22. Liu X, Lam E: Two binding sites for the plant transcription factor ASF-1 can respond to auxin treatments in transgenic tobacco. J Biol Chem 269: 668–675 (1994).
23. Medberry SL, Olszewski NE: Identification of *cis*-elements involved in Commelina yellow mottle virus promoter activity. Plant J 3: 619–626 (1993).
24. Murashige T, Skoog F: A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol Plant 15: 473–497 (1962).
25. Mushegian AR, Shepherd RJ: Genetic elements of plant viruses as tools for genetic engineering. Microbiol Rev 59: 548–578 (1995).
26. Odell JT, Nagy F, Chua NH: Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313: 810–812 (1985).
27. Olszewski N, Hagen G, Guilfoyle TJ: A transcriptionally active, covalently closed minichromosome of cauliflower mosaic virus DNA isolated from infected turnip leaves. Cell 29: 395–402 (1982).
28. Ow DW, Wood HV, DeLuca M, De Wet JR, Helinski DR, Howell SH.: Transient and stable expression of the firefly luciferase gene in plant cells and transgenic plants. Science 234: 856–859 (1986).
29. Park YD, Papp I, Moscone EA, Iglesias VA, Vaucheret H, Matzke AJ, Matzke MA: Gene silencing mediated by promoter homology occurs at the level of transcription and results

in meiotically heritable alterations in methylation and gene activity. Plant J 9: 183–194 (1996).
30. Pfeiffer P, Hohn T: Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. Cell 33: 7681–789 (1983).
31. Qin X, Holuigue L, Horvath DM, Chua NH: Immediate early transcription activation by salicylic acid via the cauliflower mosaic virus as-1 element. Plant Cell 6: 863–874 (1994).
32. Regad F, Herve C, Marinx O, Bergounioux C, Tremousaygue D, Lescure B: The tef1 box, a ubiquitous cis-acting element involved in the activation of plant genes that are highly expressed in cycling cells. Mol Gen Genet 248: 703–711 (1995).
33. Roest S, Gilissen JW: Regeneration from protoplasts – a supplementary literature review. Acta Botanica Neerlandica 42: 1–25 (1993).
34. Rothnie HM, Chapdelaine Y, Hohn T: Pararetroviruses and retroviruses: a comparative review of viral structure and gene expression strategies. Adv Virus Res 44: 1–67 (1994).
35. Sanger M, Daubert S, Goodman RM: Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous 35S promoter from cauliflower mosaic virus and the regulated mannopine synthase promoter. Plant Mol Biol 14: 433–413 (1990).
36. Teakle GR, Kay SA: The GATA-binding protein CGF-1 is closely related to GT-1. Plant Mol Biol 29: 1253–1266 (1995).
37. Verdaguer B, de Kochko A, Beachy RN, Fauquet C: Isolation and expression in transgenic tobacco and rice plants, of the cassava vein mosaic virus (CVMV) promoter. Plant Mol Biol 31: 1129–1139 (1996).
38. Vicente-Carbajosa J, Moose SP, Parsons RL, Schmidt RJ: A maize zinc-finger binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptionnal activator Opaque 2. Proc Natl Sci USA 94: 7685–7690 (1997).
39. Watanabe Y, Meshi T, Okada T: Infection of tobacco protoplasts with in vitro transcribed tobacco mosaic virus RNA using an improved electroporation method. FEBS Lett 219: 65–69 (1987).
40. Williamson JD, Hirschwyncott ME, Larkins BA, Gelvin SB: Differential accumulation of a transcript driven by the CaMV 35S promoter in transgenic tobacco. Plant Physiol 90: 1570–1576 (1989).
41. Xiang C, Miao ZH, Lam E: Coordinated activation of as-1-type elements and a tobacco glutathione *S*-transferase gene by auxins, salicylic acid, methyl-jasmonate and hydrogen peroxide. Plant Mol Biol 32: 415–426 (1996).
42. Yin Y, Beachy RN: The regulatory regions of the rice tungro bacilliform virus promoter and interacting nuclear factors in rice (*Oryza sativa* L.). Plant J: 969–980 (1995).
43. Yin Y, Chen L, Beachy RN: Promoter elements required for phloem-specific gene expression of the RTBV promoter in rice. Plant J, in press.
44. Zhang B, Singh KB: ocs element promoter sequences are activated by auxin and salicylic acid in Arabidopsis. Proc Natl Acad Sci USA 91: 2507–2511 (1994).